ADAPTIVE DISTANCE SAMPLING

John Pollard

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

2002

Adaptive distance sampling

John Pollard

Thesis submitted for the degree of

DOCTOR OF PHILOSPHY

In the School of Mathematics and Statistics,

UNIVERSITY OF ST ANDREWS.

August, 2002.

Th E180

Declarations

1. I, John Pollard, hereby certify that this thesis, which is approximately 55,000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous applications for a higher degree.

   date 5/8/2002 signature of candidate ....

2. I was admitted as a research student in April 1995 and as a candidate for the degree of PhD in September 1996; the higher study for which this is a record was carried out in the University of St Andrews between 1995 and 2002.

   date 5/8/2002 signature of candidate ....

3. I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

   date 5/8/02 signature of supervisor ....

4. In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any *bona fide* library or research worker.

   date 5/8/2002 signature of candidate ....

Abstract

We investigate mechanisms to improve efficiency for line and point transect surveys of clustered populations by combining the distance methods with adaptive sampling. In adaptive sampling, survey effort is increased when areas of high animal density are located, thereby increasing the number of observations.

We begin by building on existing adaptive sampling techniques, to create both point and line transect adaptive estimators, these are then extended to allow the inclusion of covariates in the detection function estimator. However, the methods are limited, as the total effort required cannot be forecast at the start of a survey, and so a new fixed total effort adaptive approach is developed. A key difference in the new method is that it does not require the calculation of the inclusion probabilities typically used by existing adaptive estimators. The fixed effort method is primarily aimed at line transect sampling, but point transect derivations are also provided.

We evaluate the new methodology by computer simulation, and report on surveys of harbour porpoise in the Gulf of Maine, in which the approach was compared with conventional line transect sampling. Line transect simulation results for a clustered population showed up to a 6% improvement in the adaptive density variance estimate over the conventional, whilst when there was no clustering the adaptive estimate was 1% less efficient than the conventional. For the harbour porpoise survey, the adaptive density estimate *cv*s showed improvements of 8% for individual porpoise density and 14% for school density over the conventional estimates.

The primary benefit of the fixed effort method is the potential to improve survey coverage, allowing a survey to complete within a fixed time and effort; an important feature if expensive survey resources are involved, such as an aircraft, crew and observers.

Acknowledgements

First I must thank Liz, my wife, for her support, encouragement and understanding over the last 7 years, without her help I could not have done this. Also my two children Oliver and Anna for their patience while I researched this thesis.

Steve Buckland my supervisor, for his support, guidance and insight throughout this work, and his belief that I could complete a PhD through part-time study.

Phil Hammond my second supervisor, for his advice and willingness to proof read.

To the people of Northwest Fisheries Science Center, Woods Hole, especially Debi Palka, for their work on the experimental harbour porpoise survey, and the crew and observers of the Abel-J who carried out the survey.

To everyone in RUWPA, past and present, who have always made me welcome within the group, including, David Borchers, Samantha Stringberg, Sharon Hedley, Simon Wood, Louise Burt, Liz Clarke, Fiona Thompson, Phil le Feuvre, Rhona Rodger and Cathy Brown. Finally, my thanks to Len Thomas for his freely given advice and assistance.

Contents

Chapter 1

Introduction

Wildlife abundance estimation is becoming increasingly important as the habitats and reserves of many species are diminished. At the same time the resources to assess these populations are typically limited and so it is desirable to maximise the effectiveness of any surveys performed.

Many wildlife populations occur in loose spatial clusters or aggregations and if the number of aggregations is small, then sample size may be inadequate for reliable estimation, and precision may be poor. With spatially clustered populations, it is therefore attractive to focus potentially expensive surveying resource on the spatial clusters, by increasing sampling in areas of higher detection. Hence in recent years *adaptive sampling* has been promoted as a method suited to clustered populations. Adaptive sampling adds surveying effort when the survey adapts and thus, unlike many basic sampling estimators, the analysis methods do not assume the survey effort is randomly allocated.

Distance sampling (see for example, Buckland *et al.* 2001, Buckland *et al.* 2002, Burnham *et al.* 1980, Thomas *et al.* 2002) is widely used to estimate animal abundance, particularly in the form of *line* and *point transect* sampling. Line and point transects do not require all animals within a sampled area to be detected, but instead model the probability of detection, based on the distance of the detected animals from the observer. This modelled *detection function* is then used to scale the number of observations to account for the animals that were not detected.

This thesis combines adaptive and distance methods, building on the ability of line and point transect sampling to account for imperfect detectability and adaptive sampling to improve estimator efficiency for clustered populations. The adaptive sampling is based on the adaptive cluster sampling methods of Thompson, Seber and Ramsey (Thompson, 1992; Thompson and Seber, 1996; Thompson, Ramsey and

Seber, 1992) who have done much to develop this area over the last fifteen or so years.

Whilst the initial premise is to improve estimator precision a new density estimator is developed which can also optimise survey coverage, allowing a survey to complete within a fixed total sampling effort. It does this through the efficient allocation of surveying resources and is able to compensate for lost surveying time due to bad weather or other external factors. Although precision improvements can be achieved, this is highly dependent on the underlying spatial clustering of the population, and the true benefit of the methods is likely to be the improved coverage.

1.1 Thesis Outline

Chapter 1 provides a brief overview of the adaptive cluster sampling methods of Thompson, Seber and Ramsey and also of distance sampling. Hereafter, rather than reference all three names we typically refer to just Thompson, for example, *Thompson's methods*. For brevity, we also refer to the methods as *adaptive sampling* or *adaptive methods*, and omit the complete adaptive cluster sampling title. It is expected that readers will be more familiar with distance sampling than Thompson's methods, and so the adaptive sampling is also illustrated with a brief example. We complete the chapter by introducing Thompson's basic adaptive estimators used to build the estimators of Chapters 2 and 3.

In Chapter 2 we begin by combining Thompson's adaptive methods with point transect sampling, and refer to this simply as *adaptive point transect sampling*. Four basic estimators are developed and then the methods are extended to allow the inclusion of covariates in the detection function estimation, following the approaches of Marques (2001) and Borchers (1996). A basic simulation is performed and the pros and cons of various patterns for the additional adaptive point transects explored. We close the chapter with a general discussion which also considers how the methods may be applied in the field.

Chapter 3 builds on the point transect work and develops the corresponding estimators for *adaptive line transect surveys*, again based on Thompson's methods. It follows a similar format to Chapter 2, but concentrates on the changes required for line transects. With Thompson's methods, the total surveying effort is unknown at the start of the survey. Thus we also discuss approaches for keeping the total effort within reasonable limits.

In response to the unknown total survey effort, Chapter 4 develops a new estimator, where the survey can be completed using a fixed total effort, termed the *PB method.* First a line transect estimator is developed and tested through simulations. As with the Thompson-based adaptive point and line transect estimators, the approach is extended to allow the inclusion of covariates in the detection function estimator. A point transect estimator is also derived. Much of this chapter is drawn from Pollard and Buckland (1997) and Pollard, Palka and Buckland (in press).

Chapter 5 applies the new fixed total effort estimator to an experimental harbour porpoise line transect survey. The experiment compares adaptive and conventional surveys run over the same transects in similar conditions. We explain how the survey configuration was chosen, consider the field procedures, discuss the analysis and review the results. The analysis and results of this survey were originally reported in Palka and Pollard (1999).

We close with a discussion comparing the various methods and review both the benefits and drawbacks of the adaptive procedures developed. We also consider the combination of conventional and adaptive survey methods, so that conventional estimates can be still be extracted and thus allow comparison with previous survey results.

A computer program has been developed as part of the thesis, to enable the simulation of populations, and of conventional and adaptive surveys. A brief overview of this is provided in Appendix A. Appendix B provides detailed summaries and examples of the fixed effort survey simulations described in Chapter 4.

1.1.1 Conventions

Some of the terminology used can have alternative meanings, particularly between the two approaches of adaptive and distance sampling. In this section we clarify a number of the terms to avoid confusion.

The thesis considers wildlife abundance, and thus the items in the populations sampled are typically referred to as *animals*. However the methods are not restricted to these alone.

Thompson's methods refer to the combined adaptive cluster sampling developments of Thompson, Seber and Ramsey, and *adaptive sampling* or *Thompson's adaptive methods* is used in place of adaptive cluster sampling. Thompson developed four core estimators for adaptive sampling (see section 1.3.1) and these are called the *Thompson-based estimators*, or the *basic adaptive estimators*. Thompson uses a condition which, if met, triggers adaptive sampling behaviour. This is often called the *trigger condition*, *adaptive trigger* or just *trigger*.

Within the context of this thesis, distance sampling refers to either line or point transect sampling, or some variation on these such as trapping webs. Standard distance sampling is named *conventional*, as in conventional line transect sampling or conventional point transect sampling.

When animals are detected by an observer this is called a *sighting*, a *detection* or an *observation*. In the case of populations where animals occur in groups, a single observation is known as an *object*, an *animal group* or *group*, and in some contexts, such as marine mammal surveys, as a *school*. Thus the *group size* or *school size* is the number of *individual animals* in a single observation. Within conventional distance sampling, a *cluster* typically refers to a group of animals that form a single observation. Within this thesis, a cluster is a spatial aggregation of animals or animal groups from which a number of different observations may be made by the observer. The term cluster is also used in the description of Thompson's methods to refer to a collection of associated sampling units. *Truncation distance* and *truncation half-width* both refer to the distance from the point or line beyond which observations are not used for the estimates.

The *survey region* is the area of interest for which abundance is being estimated, and the *population* the animals contained within it. Thompson's methods overlay the survey region with a grid of *sampling units*, which we also refer to as *grid units* or *units*. When the method selects a unit for sampling, then depending on the type of survey either a line transect or point transect sample is performed within the unit. We refer to the line transect *sampling strip* and the point transect *sampling plot* or *plot*, as the areas sampled from the lines or points up to some truncation distance. The *surveyed area* represents the total area of all sampling strips or plots for every unit in the grid, whether sampled or not. This may be more or less than the area of the survey region, depending on whether sampling strips/plots in adjoining units overlap. To accommodate this, when producing estimates for the survey region, the estimates relating to the surveyed area have to be adjusted according to ratio of its area to the survey region area.

Where possible, notation is maintained in accord with both distance sampling and Thompson's methods. However this has not always been possible, due to the number of parameters in use, and there have been some clashes. In such cases, if feasible, context is maintained by using a similar letter from an alternative alphabet or at least a letter that represents the parameter. Thus for example in the adaptive point transect chapter, w was already used in a number of the Thompson estimators and so the truncation radius is represented by R, whilst in the adaptive line transect chapter the truncation half-width is represented by W.

1.2 Overview of Distance Sampling

Distance sampling is an extension of quadrat sampling methods, such as *strip transect sampling* and *point counts*. In strip quadrat sampling an observer travels down the centreline of long narrow strips, counting all objects that lie within the strip. Similarly in point counts, the observer is located at a point and counts all objects that lie within a circle of fixed radius. In both cases the user is required to observe all objects and so the truncation distance, beyond which observations are not included, has to be kept small. *Line* and *point transect sampling* are the core methods of distance sampling and extend the quadrat sampling methods by relaxing the requirement to detect all objects. Instead the observer records the distance to each observation, and the sample distances are used to model the probability of detecting an object based on its distance from the line or point, referred to as the *detection function*. The detection function along with the sample size can then be used to estimate the actual number of animals in the sampled strip or circle, which is then used to provide an overall estimate of abundance in the survey region. The basic estimators assume: that all animals are detected on the line or point; that the probability of detecting an object decreases as the distance from the observer increases; that the distances to the objects are accurately recorded; that objects are detected at their initial position and there is no movement in response to the observer; and that the lines or points have been placed randomly with respect to the distribution of objects. Approaches are available to relax these assumptions but for this introduction we concentrate on the basic methods.

Free software is available for the analysis of conventional distance sampling (Thomas *et al.*, 2002; Laake *et al.*, 1996) and a comprehensive introduction to the methods and their application is provided by Buckland *et al.* (2001).

1.2.1 Line Transect Sampling

In line transect sampling a number of randomly placed lines are traversed by the observer, and the perpendicular distances to all objects observed are recorded. In some cases the observer may instead record the radial distance to the object and the angle between the sight line and the trackline being followed, which is then easily converted to a perpendicular distance using basic trigonometry. Let the total length of transect surveyed be L, the number of objects observed n, and the detection function $g(x)$, where x is the distance from the line.

A probability density function (pdf) of perpendicular distances to detected objects $f(x)$, is fitted to the distance data so that the *effective half-width* μ can be estimated. If detection on the line is certain then it can be shown that $\mu = 1/f(0)$, where $f(0)$ is the pdf of the detection function evaluated at $x = 0$ (Buckland *et al.*, 2001: p53). The effective half-width relates to the width of a strip of total length L (the half-width being the width of strip on either side of the trackline) so that had all objects been detected within that strip, we would expect to detect $E(n)$ observations. Thus such a strip would have area $2\mu L$ and an estimate of the object density is given by

$$\hat{D}_{LT} = \frac{E(n)}{2\hat{\mu}L}$$

If each observation is of a group rather than an individual animal, then the density of individual animals is given by

$$\hat{D}_{LT} = \frac{E(n)\cdot\hat{E}(s)}{2\hat{\mu}L}$$

where

$\hat{E}(s)$ is an estimate of the expected mean group size in the population

1.2.2 Point Transect Sampling

In point transect sampling, the observer is located at a point and records the radial distances to observations. In this case let the total number of points sampled be k, the total number of objects observed n, and the detection function $g(r)$, where r is the radial distance from the point to an object. For point transects we want to estimate the effective radius ρ, which is the radius of a circle such that if all objects were detected from each of k points, we would expect to detect $E(n)$ observations. Thus the effective area for a point is $\pi\rho^2$ and the total effective area across all the sampled points is $k\pi\rho^2$. An estimate of the object density is given by

$$\hat{D}_{PT} = \frac{E(n)}{k\pi\rho^2}$$

If the observations relate to groups of animals, then the density estimate can be converted to a density of animals by multiplying by an estimate of the expected mean group size.

Estimating the effective radius for point transects is slightly more complex than estimating the effective strip half-width for line transects. This is because with line transects, the area of the sampling strip increases linearly as the width increases, whilst for point transects, the area of the sampling circle increases as the square of the radius. Thus in line transects the pdf of the distances has the same shape as the detection function, but this is not the case for point transects. For points, if detection on the point is certain, then it can be shown that $\rho = \sqrt{2/h(0)}$, where $h(0)$ is the derivative of the pdf evaluated at $r = 0$ (Buckland *et al.*, 2001: p55).

1.3 Overview of Adaptive Cluster Sampling

To give an insight into Thompson's methods, we provide a brief overview of adaptive cluster sampling, hereafter referred to as adaptive sampling. The approach is aimed at the sampling of rare, but spatially clustered populations, and works by sampling additional units, above the initial sample, when the variable of interest for a sampled unit meets some trigger condition. In our case, this is typically when the count of animals exceeds some preset limit.

The basic adaptive process for Thompson's methods operates as follows. A number of *initial units* are selected at random and sampled. If the number of observations in a sampled unit satisfies some *condition* (also termed here the trigger condition, or trigger), then units in the neighbourhood of the triggering unit are also sampled. The neighbourhood defines a symmetric pattern of units and its layout is part of the survey design. If any of the adaptive units in the neighbourhood meet the condition, then the neighbourhood of each of these units is also sampled. The process repeats until no newly sampled units meet the condition.

The combination of an initial unit and its associated adaptive units is termed a *cluster*. Within the cluster, any units which do not meet the condition are termed *edge units*, whilst any units which meet the condition form a *network*. Any initial unit that does not meet the condition is also a network, consisting of a single unit. The neighbourhood can be defined as any pattern surrounding the sampled unit, and may not be contiguous. However it must have the property that the sampling of any unit within a network will then also sample all other units within the same network.

The final sample will therefore consist of a network for each of the initially sampled units. However it is possible that the networks from two or more separate initial units may merge into one larger network.

To illustrate this consider a simple example, where the survey region has been overlaid with a grid of sampling units and nine initial sampling units have been randomly selected (Figure 1.1). If a unit is sampled then all animals within that unit are detected.

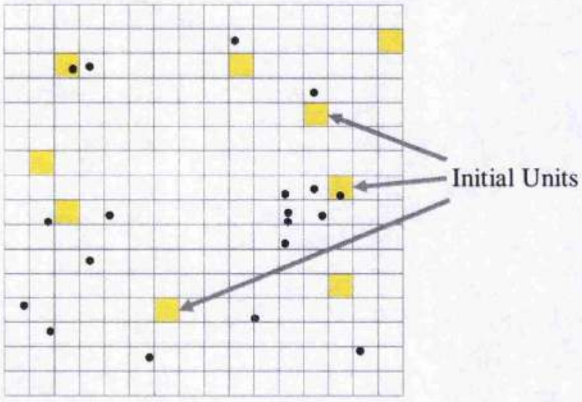


Figure 1.1: A grid of sampling units is overlaid on the survey region. Initial sampling units have been selected and are shown as shaded (yellow). Animals in the population are shown as black dots.

In this example the trigger condition is defined as one or more observations in a sampling unit, and adaptive units are added on the four adjoining edges, above, below and to the left and right, of any unit that meets the condition. Two initial units meet the condition and so adaptive units are added to these units (Figure 1.2).

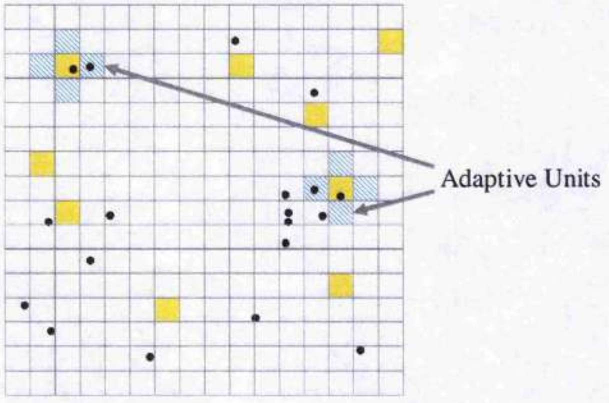


Figure 1.2: Adaptive units are added in the four adjacent units above, below and to the left and right of any initial units that meet the trigger condition of one or more observations. Adaptive units are shown with (blue) cross-hatching.

The process continues for each additional adaptive unit that triggers the condition, and stops when no new units meet the condition. Each initial unit together with its associated adaptive units forms a cluster. As shown in Figure 1.3, a total of twenty two adaptive units are sampled.

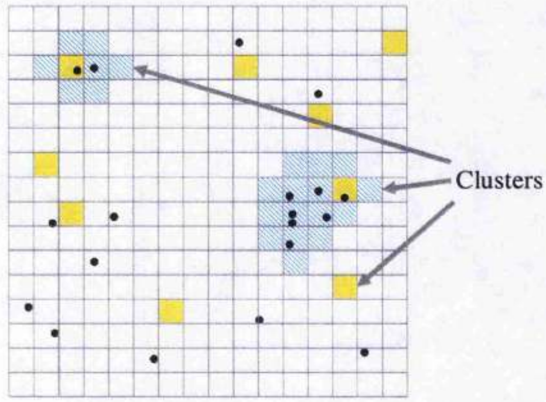


Figure 1.3: The adaptive process continues for any adaptive units that meet the trigger condition, until no new units meet the condition. Each initial sampling unit together with any associated adaptive units forms a cluster.

Each cluster consists of a network and edge units. There are nine networks in the sample, seven of which are single initial units which did not meet the trigger condition. Of the two initial units that did meet the condition, one forms a network of size two with six edge units in its cluster; whilst the other is a network of six units with ten edge units in its cluster (Figure 1.4). Had the adaptive units spread sufficiently so that the networks of two (or more) initial units joined, then these would have formed a single network.

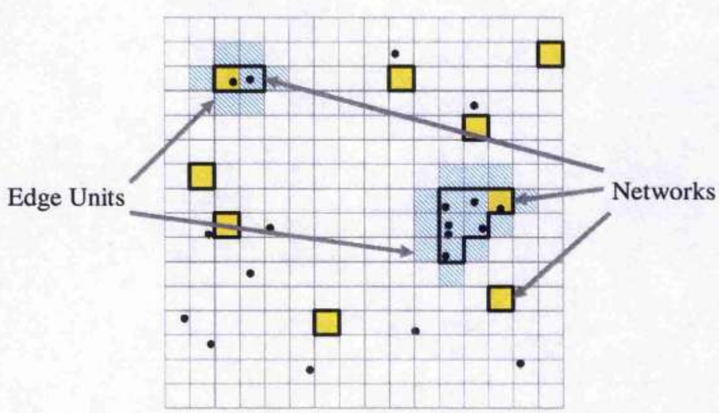


Figure 1.4: Initial units that do not meet the trigger condition form networks of size one. Initial units which meet the trigger condition form networks with any adaptive units that also meet the condition. These are surrounded by edge units, which are the adaptive units within the cluster which do not meet the condition. The networks are shown as enclosed in a thick black line.

1.3.1 Adaptive Estimators

Thompson's estimators are typically *design-unbiased* (Thompson, 1992: p94; Thompson and Seber, 1996: p14). In a *design-based* abundance estimator, the values of the variable of interest, for example the count of animals in each quadrat, are considered fixed and the selection probabilities introduced by the design are used to estimate the abundance and associated variance, etc. Thus a design-unbiased estimator is unbiased whatever the underlying population.

The Thompson-based estimators are derived from the unequal probability sampling estimators of Hansen-Hurwitz (Hansen and Hurwitz, 1943) and Horvitz-Thompson (Horvitz and Thompson, 1952). We now describe these two estimators with respect to a simple random sample of units (quadrats) from a survey region with the variable of interest being the count of animals made in each unit sampled.

The Hansen-Hurwitz is an unbiased estimator for sampling with replacement. It considers the draw-by-draw selection probabilities for each unit in the sample and these probabilities are used to weight the sample size for that unit in the sample, so that the sum of these weighted values provides an estimator of the population total. If a unit is drawn twice, then its count is used twice in the summation. The Hansen-Hurwitz estimator for the population total and its variance are

$$\hat{\tau}_{HH} = \frac{1}{k}\sum_{i=1}^{k}\frac{y_i}{p_i} \qquad \text{and} \qquad \hat{V}(\hat{\tau}_{HH}) = \frac{1}{k(k-1)}\sum_{i=1}^{k} p_i\left(\frac{y_i}{p_i} - \hat{\tau}_{HH}\right)^2$$

where

k	is the number of units in the sample
y_i	is count of interest in for the i^{th} unit
p_i	is the probability of selecting the i^{th} unit in a draw

The Horvitz-Thompson estimator considers the probability of including any unit in the sample. It only uses the distinct units in the sample, so that if a unit is drawn multiple times it is only used once in the estimator. The Horvitz-Thompson estimator is unbiased for sampling both with and without replacement. Its estimate of the population total and its variance are:

$$\hat{\tau}_{HT} = \sum_{i=1}^{v}\frac{y_i}{\pi_i} \quad \text{and} \quad \hat{V}(\hat{\tau}_{HT}) = \sum_{i=1}^{v}\left(\frac{1-\pi_i}{\pi_i}\right)y_i + 2\sum_{i=1}^{v}\sum_{j\neq i}\left(\frac{\pi_{ij}-\pi_i\pi_j}{\pi_i\pi_j\pi_{ij}}\right)y_i y_j$$

where

v	is the number of distinct units in the sample

y_i is value of interest for the i^{th} unit

π_i is the probability of including the i^{th} unit in the sample

π_{ij} is the probability of including both the i^{th} and j^{th} units in the sample

Rather than considering the count of animals in each quadrat, Thompson's adaptive estimators use the count of animals in each network and consider the probability of the networks being included. When the trigger condition is a single detection, there are no animals detected in edge units, however if the trigger condition is greater than one, then the count of animals in any edge units are not included in the sample when producing estimates.

We complete the introduction by outlining the basic adaptive sampling estimators for Thompson's methods. We consider two categories of adaptive sampling design and for each design Thompson has both a Hansen-Hurwitz and a Horvitz-Thompson-based estimator, making a total of four basic estimators.

The two survey designs are a Random Initial Sample (RIS) and a Systematic Initial Sample (SIS). As the name infers, a RIS design survey has the initial units selected at random, and adaptive units are added to any initial units which meet the trigger condition. The RIS design estimators for this are described further in Thompson (1990). A SIS design consists of primary units and secondary units. The secondary units are systematically arranged to form the primary units. The primary units are however still randomly selected. In this case adaptive units are added to any secondary units that meet the trigger condition. More detailed descriptions of the SIS design estimators are found in Thompson (1991). In Chapters 2 and 3 the two designs are explained in more detail with specific reference to point and line transect surveys.

The notation used here is typically directly from Thompson (1992) or Thompson and Seber (1996). In later chapters there is often a clash of notation, and it has had to be adjusted on a case by case basis.

RIS Design Estimators

RIS Design: Hansen-Hurwitz-based Estimator

Thompson developed an estimator from the Hansen-Hurwitz estimator for sampling with and without replacement (Thompson, 1992: p271). The estimate of the mean number of objects per unit is

$$\hat{\mu}_1 = \frac{1}{k}\sum_{i=1}^{k} w_i \qquad (1.1)$$

where

$$w_i = \frac{1}{m_i}\sum_{j \in \psi_i} y_j$$

k	is the number of units in the initial sample
y_j	is the y value (value of interest) for the j^{th} unit in network ψ_i
ψ_i	is the network which includes the i^{th} initial unit
m_i	is the number of units in network ψ_i

and the corresponding variance estimator, assuming the initial sample is selected without replacement, is

$$\hat{\operatorname{var}}[\hat{\mu}_1] = \frac{(K-k)}{Kk(k-1)}\sum_{i=1}^{k}(w_i - \hat{\mu}_1)^2 \qquad (1.2)$$

where

K	is the total number of units in the survey region

RIS Design: Horvitz-Thompson-based Estimator

Thompson has also developed an estimator from the Horvitz-Thompson estimator for sampling with or without replacement (Thompson, 1992: p273). The without replacement estimator of the mean number of objects per unit is

$$\hat{\mu}_2 = \frac{1}{K}\sum_{i=1}^{\nu}\frac{y_i}{\alpha_i} \qquad (1.3)$$

where

$$\alpha_i = 1 - \binom{K-t_i}{k} \Big/ \binom{K}{k}$$

ν	is the number of distinct networks in the sample
k	is the number of units in the initial sample

y_i	is the sum of the y values (value of interest) for the i[th] network
α_i	is the probability that the i[th] network is included in the sample
K	is the total number of units in the survey region
t_i	is the number of units in the i[th] network.

and the corresponding variance estimator is

$$\hat{\text{var}}[\hat{\mu}_2] = \frac{1}{K^2} \sum_{i=1}^{\nu} \sum_{h=1}^{\nu} \frac{y_i y_h (\alpha_{ih} - \alpha_i \alpha_h)}{(\alpha_i \alpha_h \alpha_{ih})} \tag{1.4}$$

where

$$\alpha_{ih} = 1 - \left\{ \binom{K - t_i}{k} + \binom{K - t_h}{k} - \binom{K - t_i - t_h}{k} \right\} \Big/ \binom{K}{k}$$

SIS Design Estimators

SIS Design: Hansen-Hurwitz-based Estimator

From Thompson (1992: p293), for a systematic or strip adaptive survey, an estimate of the mean number of objects per unit is given by

$$\hat{\mu}_3 = \frac{1}{r} \sum_{h=1}^{r} w_h$$

where

$$w_h = \frac{1}{M} \sum_{i=1}^{K_h} \frac{y_i}{t_i} \tag{1.5}$$

r	is the number of primary units in the sample
M	is the number of secondary units in each primary unit
K_h	is the number of networks that intersect the h[th] primary unit
y_i	is the sum of the y values (value of interest) in the i[th] network
t_i	is the number of primary units that intersect the i[th] network

An unbiased estimate of the variance is given by

$$\hat{\text{var}}[\hat{\mu}_3] = \frac{s_W^2}{r}\left(1 - \frac{r}{R}\right) \tag{1.6}$$

where

$$s_W^2 = \frac{1}{r-1} \sum_{h=1}^{r} (w_h - \hat{\mu}_3)^2$$

R	is the number of primary units in the survey region

This is based on the initial intersection probabilities and is related to the Hansen-Hurwitz estimator.

SIS Design: Horvitz-Thompson-based Estimator

From Thompson (1992: p295), for a systematic or strip adaptive survey, an estimate of the mean number of objects per unit is given by

$$\hat{\mu}_4 = \frac{1}{MR}\sum_{i=1}^{v}\frac{y_i}{\alpha_i} \tag{1.7}$$

where

$$\alpha_i = 1 - \binom{R-t_i}{r} \Big/ \binom{R}{r}$$

M	is the number of secondary units in each primary unit
R	is the number of primary units in the survey region
y_i	is the sum of the y values (value of interest) in the i^{th} network
α_i	is the probability the i^{th} network is included in the sample
t_i	is the number of primary units that intersect the i^{th} network
r	is the number of primary units in the sample
v	is the number of distinct networks in the sample

The corresponding variance estimator is

$$\hat{\mathrm{var}}[\hat{\mu}_4] = \frac{1}{M^2R^2}\sum_{i=1}^{v}\sum_{h=1}^{v}\frac{y_i y_h(\alpha_{ih} - \alpha_i\alpha_h)}{(\alpha_i\alpha_h\alpha_{ih})} \tag{1.8}$$

where

$$\alpha_{ih} = 1 - \left\{\binom{R-t_i}{r} + \binom{R-t_h}{r} - \binom{R-t_i-t_h+t_{ih}}{r}\right\} \Big/ \binom{R}{r}$$

and

t_{ih}	is the number of primary units that intersect networks i and h

Chapter 2

Adaptive Point Transect Sampling

2.1 Introduction

This chapter explores the application of Thompson's adaptive methods (see, for example, Thompson, 1992; Thompson and Seber, 1996) to point transect surveys.

Point transect sampling (Buckland *et al.*, 2001), also known as variable circular plots, is most commonly used in ornithology. Extensions to the method include trapping webs (Anderson *et al.*, 1983; Buckland *et al.*, 2001; Parmenter *et al.*, 2002) and cue counting (Hiby, 1982; Hiby, 1985; Hiby and Hammond, 1985; Buckland *et al.*, 2001).

The typical approach is to define a series of points in the survey region. These points can be located either randomly; on lines located randomly or systematically; or on a grid randomly located on the area. At each point the observer records the distance to any animals seen there. All animals close to the point must be detected whilst as the distance increases it is expected that the proportion of animals detected will decrease. The detection distances are used to estimate the detection function, which is in turn used to estimate the effective area. That is the area of a circle at which, assuming all animals to be detected, would produce the same count of detections as was actually recorded. Alternatively this can be considered as the area of a circle such that as many animals are seen outside this circle as are missed inside it. Assuming single animals for each observation, the density estimate is then simply given by the mean number of observations per point divided by the effective area.

Point transects have a number of advantages over line transects, which encourage their use for surveying birds. The observer is located at a point and so can concentrate on detection without also having to negotiate a transect through potentially difficult terrain. Observations are made from a point rather than along a line so it is more suited to surveying patchy habitats and the observer can take the most direct or easiest route to and from each point. In addition markers can be positioned at set distances, making the observer's task easier, and only the distance and not the angle is required to be estimated by the observer.

A point transect survey typically has at least 20 points, but there may need to be more to get sufficient observations for reliable estimation of the detection function. For rare species this may require hundreds of points. To provide acceptable levels of precision the survey should aim for a minimum of around 75-100 detections of the prime species being surveyed (Buckland *et al.*, 2001: p240).

Within this chapter we combine adaptive with point transect sampling by treating each point as a sampling unit. Thompson's adaptive estimates are used to obtain estimates of the mean number of detections per unit and the expected group size whilst sightings data are pooled across all sightings to improve estimation of the detection function. We start by developing combined adaptive and point transect sampling estimators for two basic survey designs; a random and a systematic initial sample. These terms are explained in the following section, Survey Designs. For each design both a Hansen-Hurwitz and a Horvitz-Thompson-based estimator are developed, making a total of four basic estimators. Many surveys will use systematically positioned initial points, however it is widely accepted that if the grid is randomly located on the survey region then the points can be considered random. Thus later in the chapter we focus on a Horvitz-Thompson-based estimator with a random initial sample.

The estimators include a number of subscripted variables which can at first appear confusing. For clarity a simple example, with full working, is presented for each of the four basic estimators considered.

A disadvantage of Thompson's methods is that the total effort required is not known at the beginning, making it difficult to identify the total cost and time to complete the survey. However it may be possible to estimate total effort based on some scale function utilising previous survey data. Approaches to deal with this are discussed in Chapter 3, Adaptive Line Transect Sampling.

The observers will need to move between points, referred to here as *off-effort*, where off-effort encompasses both the time taken and any resource used in the travel (e.g. a vehicle). By the very nature of point transect sampling, the additional points sampled in the adaptive neighbourhoods will introduce an amount of off-effort travel. However for a point transect survey this is likely to be small, when compared to the off-effort travel between the points on a conventional survey. Thus the total off-effort travel for an adaptive survey will potentially be less than that for a conventional point transect survey, with an equivalent number of points surveyed. This is akin to the gain from cluster sampling but in this case the sampled clusters are expected to occur where animals are known to be present.

2.1.1 Survey Designs

In this chapter we consider two types of survey design, one classified as a Random Initial Sample (RIS) and the other as a Systematic Initial Sample (SIS). The terms refer to the selection of the initial survey points; adaptive points are then added according to the adaptive neighbourhood pattern in use. These two terms are not immediately intuitive and warrant further explanation.

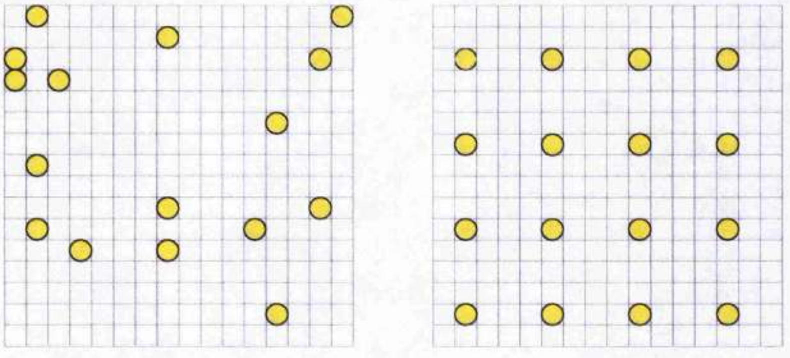


Figure 2.1: RIS point transect survey designs, with the sampling points shown as solid (yellow) circles. In the left hand design points have been selected at random on the grid. The right hand design is a systematic grid of points, and (given a random start location) the points are assumed to be independently located.

If the points are selected at random from the survey region, then this is clearly a RIS design. Typical systematic arrangements of points transects are lines of points, with the lines either randomly or systematically spaced. If the spacing between points on each line is similar to the space between lines, we obtain a systematic grid of points. In distance sampling, if the systematic grid is randomly located, it is commonplace to consider the points as random and hence this is considered a RIS design, with the result that variance is slightly overestimated. If the interline spacing is significantly different from the point spacing on lines, it should be treated as a systematic sample and thus a SIS design. Examples of RIS designs are shown in Figure 2.1 and SIS designs in Figure 2.2.

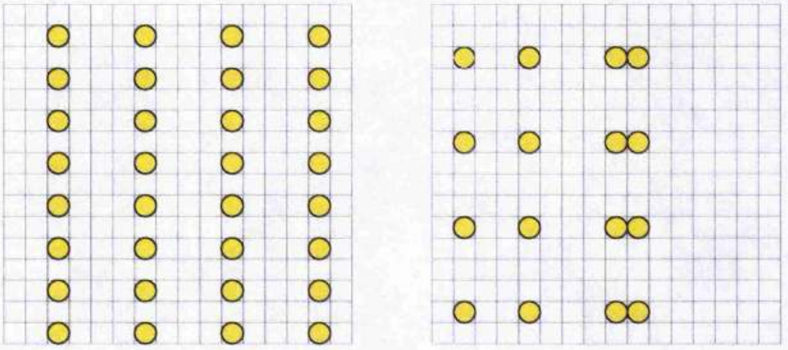


Figure 2.2: SIS point transect survey designs, with the sampling points shown as solid (yellow) circles. In the left hand design, points are systematically located in lines (the primary units), which are then systematically distributed over the survey region. In the right hand design points are also systematically distributed in lines, which in this case are then randomly (in the horizontal plane) located on the survey region.

2.2 Theory

In this section we develop the adaptive point transect estimators. Point transect sampling is commonly used for songbirds where animals are typically recorded as individuals. However throughout this chapter we follow a generic approach and allow for observations to be a single or multiple animals. Should animals be recorded as individuals then the number of animals per observation is set to 1 and all estimators will function correctly.

2.2.1 Point Transect Sampling Basic Formulae

From Buckland *et al* (2001: p55) the density of a population from point transect sampling is given by

$$D = \frac{E(n) \cdot h(0) \cdot E(s)}{2\pi k}$$

where

$E(n)$	is the expected number of animal groups in the sample
$h(0)$	is the derivative of the probability density function $f(r)$ evaluated at r=0
$E(s)$	is the expected group size for the population
k	is the number of points

Replacing parameters by their estimators, an estimate of the density is given by

$$\hat{D} = \frac{n \cdot \hat{h}(0) \cdot \hat{E}(s)}{2\pi k} \tag{2.1}$$

Using the delta method (Seber, 1982: p5-7) the variance can be estimated by

$$\hat{\mathrm{var}}(\hat{D}) = \hat{D}^2 \cdot \left[\frac{\hat{\mathrm{var}}[n]}{n^2} + \frac{\hat{\mathrm{var}}[\hat{h}(0)]}{[\hat{h}(0)]^2} + \frac{\hat{\mathrm{var}}[\hat{E}(s)]}{[\hat{E}(s)]^2} \right] \tag{2.2}$$

2.2.2 Merging Adaptive and Point Transect Sampling

Defining the *survey region* as the area of interest and the *population* as animals contained within it, the basic approach is to overlay the survey region with a grid of units. Units are then selected according to some sampling algorithm and within each of these units a point transect survey is conducted. For each point sampled the surveyed circle, up to the truncation radius, is referred to as the *plot*, with the centre of the plot centred within the grid unit. The shape of the units and the truncation radius of the point transect are discussed in section 2.3, Grid Design; however initially it is probably easiest to visualise the units as squares, as shown in Figure 2.1 and Figure 2.2.

Thus each unit represents a potential point in the survey region, and there will be a total of K points in the survey region. We define the *surveyed area* as the area of each plot multiplied by the total number of potential points in the survey region, K. The area of the point transect plots may be larger or smaller than the area of the units. If the plot is larger then adjacent plots will overlap, whilst if it is smaller there will be gaps between adjacent plots. Thus the surveyed area can be smaller or greater than the actual area of the survey region (Figure 2.3).

Thompson's methods are used to get unbiased estimates for the key parameters of the expected number of observations (groups detected, where a group may be a single or multiple animals), and the expected group size. However we pool observations across all points to produce an $h(0)$ estimate. This assumes no heterogeneity is introduced from the adaptive process and we initially assume that the probability of detection on the point, g_0, is 1. The estimates are then fed into the point transect equations to provide a density estimate.

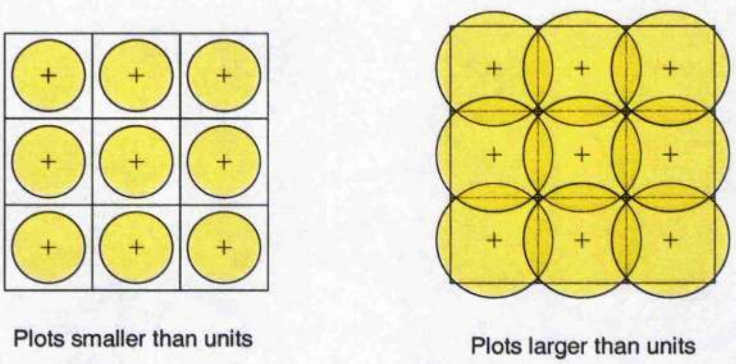


Figure 2.3: Point transect plots may be larger or smaller than the units they are located within. Thus the surveyed area is given by the total number of plots, *K*, in the survey region, multiplied by the area of each plot. Point transect plots are shown in yellow, on a grid of square units.

It should be noted that in conventional distance sampling density estimates, we use n as the number of observations in the sample. Here we use Thompson's methods to estimate $\overline{n}$, the mean number of observations per point, and $\overline{a}$, the mean number of individual animals detected per point. Note that in general an observation is a group of animals: $\overline{a} \geq \overline{n}$. We also use N to represent all animals in the survey region and N_g as the number of animal groups in the survey region.

Thus

$$N = \sum_{X=1}^{N_g} s_X$$

where

s_X is the number of animals in the X$^{\text{th}}$ observation

Later in the chapter we extend the methods to get an estimate of τ, the number of animal groups (detected and undetected) in the surveyed area, and δ, the total number of individual animals (detected and undetected) in the surveyed area. The surveyed area is a notional area and, as already stated, may be larger or smaller than the actual survey region. Thus the total number of animal groups in the survey region is obtained by appropriate scaling of τ:

$$N_g = \frac{\text{Survey Region Area}}{\text{Surveyed Area}} \cdot \tau \tag{2.3}$$

Similarly, the total number of animals in the survey region is given by

$$N = \frac{\text{Survey Region Area}}{\text{Surveyed Area}} \cdot \delta \tag{2.4}$$

Imperfect Detectability

Thompson's methods are typically founded on all animals being observed within a unit, whilst the underlying premise of distance sampling is that only a proportion of animals are observed. Thompson and Seber (1996: Chapter 9) provide approaches to deal with this, based on dividing the population estimate by the probability of detection, as is done with distance sampling. We show here that this means the underlying Thompson based adaptive estimators for the count of objects and animals remain unchanged.

From a distance sampling perspective the probability of detection is estimated and for many practical applications this can be assumed to be the same for each object in the population. So, initially, we assume the probability of detection is equal for all objects, and this is included via the pooled $h(0)$ estimate.

For an estimated equal probability of detection for all objects, Thompson and Seber (1996: p219) simply assume that the expected number of objects detected, $E(\tau_0)$, in the surveyed area, is given by the number of objects in the surveyed area, τ, multiplied by the probability of detection, p. So that

$$E(\tau_0) = \tau p$$

and thus an estimate of the number of objects in the surveyed area is given by

$$\hat{\tau} = \frac{\hat{\tau}_0}{\hat{p}}$$

and the variance is simply estimated using the delta method (Seber, 1982: p5-7), using the components from the adaptive count estimate and from the estimate of detection probability. This is analogous to the way the point transect density estimator (see equation 2.1) incorporates an estimate of the probability of detection at the point, using $h(0)$. This probability of detection is used to scale up the density estimate, accounting for missed animals. The variance is then estimated with components from: the number of observations; the detection function estimate; and the group size estimate.

Thus with an equal probability of detection (at distance x from the point) for all animals the adaptive formulae of Thompson and Seber are directly applied with the distance sampling equations.

However methods are also available to allow for unequal detection probability, so that each animal has a different probability of inclusion in the sample. See, for example, Borchers (1996), Marques (2001) and Strindberg (2001). Later in the chapter this type of approach is used to expand the framework and allow each animal to have a unique probability of detection. This can also be used to relax the requirement for all objects to be detected on the point, which would in the past have been managed by incorporating an estimate of g_0 into the divisor of the density estimate (e.g. Buckland *et al.*, 1993: p57).

Expected Number of Observations

The expected number of observations across the k_s points surveyed is $E(n)$. Thus the expected number of observations per point (the expected encounter rate), $E(\bar{n})$ is given by

$$E(\bar{n}) = \frac{E(n)}{k_s}$$

so that

$$E(n) = E(\bar{n}) \cdot k_s$$

and an estimate of the expected mean number of observations in the sample is

$$\hat{E}(n) = \hat{E}(\bar{n}) \cdot k_s$$

Expected Group Size

We estimate the expected group size by the mean observed group size, that is, as the total number of individual animals observed divided by the number of observations.

Let

a	be the total number of animals observed in the sample
a_j	be the total number of animals observed at the j[th] point
a_{jX}	be the number of animals in the X[th] observation of the j[th] point

So

$$a_j = \sum_{X=1}^{n_j} a_{jX}$$

where

n_j	is the number of detections (animal groups) at the j[th] point

and

$$a = \sum_{j=1}^{k_s} a_j$$

where

k_s	is the number of points surveyed

Let $E(a)$ be the expected total number of animals observed in the sample, had there been no adaptation, and $E(\overline{a})$ the expected mean number of animals observed per point. Then

$$E(a) = E(\overline{a}) \cdot k_s$$

so that an estimate of the expected number of animals observed in the sample, is

$$\hat{E}(a) = \hat{E}(\overline{a}) \cdot k_s$$

and an estimate of the expected group size is given by

$$\hat{E}(s) = \frac{\hat{E}(a)}{\hat{E}(n)} = \frac{\hat{E}(\overline{a}) \cdot k_s}{\hat{E}(\overline{n}) \cdot k_s} = \frac{\hat{E}(\overline{a})}{\hat{E}(\overline{n})}$$

Density Estimates

The conventional distance based density estimator (equation 2.1) divides the total number of observations for all points in the sample, by an effective sample area, where the effective area is the total area sampled but scaled to account for missed animals. However from Thompson's estimators we have an estimate of the mean number of observations per point, and thus we need to scale this by the number of points in the sample, to provide an equivalent estimator.

So substituting Thompson's estimators for the number of observations and mean group size estimator in the point transect estimators (equations 2.1 and 2.2) gives a density estimate of:

$$\hat{D} = \frac{\hat{E}(n) \cdot \hat{h}(0) \cdot \hat{E}(s)}{2\pi k_s} = \frac{\hat{E}(\bar{n}) \cdot k_s \cdot \hat{h}(0) \cdot \hat{E}(s)}{2\pi k_s} = \frac{\hat{E}(\bar{n}) \cdot \hat{h}(0) \cdot \hat{E}(s)}{2\pi}$$

where

$\hat{E}(n)$	is an estimate of the expected number of observations in the sample
$\hat{E}(\bar{n})$	is an estimate of the expected mean number of observations per point
$\hat{h}(0)$	is an estimate of the derivative of the probability density function $f(r)$ evaluated at r=0
$\hat{E}(s)$	is an estimate of the expected group size for the population.
k_s	is the number of points in the survey

and

$$\hat{\text{var}}(\hat{D}) = \hat{D}^2 \cdot \left[\frac{\hat{\text{var}}[\hat{E}(\bar{n})]}{[\hat{E}(\bar{n})]^2} + \frac{\hat{\text{var}}[\hat{h}(0)]}{[\hat{h}(0)]^2} + \frac{\hat{\text{var}}[\hat{E}(s)]}{[\hat{E}(s)]^2} \right]$$

With the group size estimated by the mean observed group size then the estimate simplifies to

$$\hat{D} = \frac{\hat{E}(\bar{n}) \cdot \hat{h}(0) \cdot \hat{E}(s)}{2\pi} = \frac{\hat{E}(\bar{n}) \cdot \hat{h}(0) \cdot \hat{E}(\bar{a})}{2\pi \cdot \hat{E}(\bar{n})} = \frac{\hat{h}(0) \cdot \hat{E}(\bar{a})}{2\pi} \tag{2.5}$$

and

$$\hat{\text{var}}(\hat{D}) = \hat{D}^2 \cdot \left[\frac{\hat{\text{var}}[\hat{h}(0)]}{[\hat{h}(0)]^2} + \frac{\hat{\text{var}}[\hat{E}(\bar{a})]}{[\hat{E}(\bar{a})]^2} \right] \tag{2.6}$$

The density of animal groups, D_g, is obtained by replacing the group size estimator, $\hat{E}(s)$ by 1 giving

$$\hat{D}_g = \frac{\hat{E}(\bar{n}) \cdot \hat{h}(0)}{2\pi} \tag{2.7}$$

and

$$\hat{\mathrm{var}}\left(\hat{D}_g\right) = D_g^{\ 2} \cdot \left[\frac{\hat{\mathrm{var}}\left[\hat{E}(\overline{n})\right]}{\left[\hat{E}(\overline{n})\right]^2} + \frac{\hat{\mathrm{var}}\left[\hat{h}(0)\right]}{\left[\hat{h}(0)\right]^2} \right] \tag{2.8}$$

2.2.3 Assumptions

The basic point transect sampling assumptions apply, although these can be weakened or removed using advanced distance sampling strategies:

(i) Probability of detection on the point is 1, or suitable methods are used to estimate g_0.

(ii) There is no size bias (the probability of detection is independent of group size).

(iii) There is no responsive movement of animals in advance of detection.

In addition, we begin with the following additional assumption in place:

(iv) Probability of detection is independent of whether or not effort is adaptive. i.e. probability of detection is only a function of distance from the point and the adaptive procedure does not induce heterogeneity in the *h(0)* estimate.

Approaches to reduce the need for assumptions (i), (ii) and (iv) are dealt with later in this chapter.

2.2.4 Adaptive Point Transect Sampling with RIS Design

We start by considering a point transect survey combined with adaptive sampling, where the initial points are selected at random within the survey region. Thompson begins with two basic estimators for this case. The first is based on the draw-by-draw probabilities, that the initial sample will intersect a unit's network and is formed from the Hansen-Hurwitz estimator. The second estimator is based on the probabilities of the initial sample intersecting networks and is developed from the Horvitz-Thompson estimator. In this and the following sections we develop estimators of the expected mean number of observations per point, $E(\overline{n})$, and the expected mean number of animals observed per point, $E(\overline{a})$.

RIS Design: Hansen-Hurwitz-based estimators

Taking Thompson's Hansen-Hurwitz-based estimator, and concentrating on sampling without replacement (equations 1.1 and 1.2), we produce unbiased estimates of the expected mean number of observations per point, $\hat{E}(\bar{n}_{HH})$, and the expected mean number of animals observed per point, $\hat{E}(\bar{a}_{HH})$.

$E(\bar{n}_{HH})$ Estimate for RIS Design

Applying equations 1.1 and 1.2 to the point transect surveys gives an estimate of expected mean number of observations per point of

$$\hat{E}(\bar{n}_{HH}) = \frac{1}{k}\sum_{i=1}^{k} w_i \tag{2.9}$$

where

$$w_i = \frac{1}{m_i}\sum_{j \in \psi_i} n_{ij}$$

k	is the number of points in the initial sample
ψ_i	is the network which includes the i[th] initial point
m_i	is the number of points in the network ψ_i
n_{ij}	is the number of observations at the j[th] point within the network ψ_i

The estimate of variance, assuming the initial sample is selected without replacement, is given by

$$\hat{V}\left(\hat{E}(\bar{n}_{HH})\right) = \frac{(K-k)}{Kk(k-1)}\sum_{i=1}^{k}\left(w_i - \hat{E}(\bar{n}_{HH})\right)^2 \tag{2.10}$$

where

K	is the number of (potential) points in the survey region

$E(\bar{a}_{HH})$ Estimate for RIS Design

Similarly an estimate of the expected mean number of animals observed per point is given by:

$$\hat{E}(\bar{a}_{HH}) = \frac{1}{k}\sum_{i=1}^{k} u_i \tag{2.11}$$

where

$$u_i = \frac{1}{m_i} \sum_{j \in \psi_i} \sum_{X=1}^{n_{ij}} a_{ijX}$$

a_{ijX} is the number of animals in the X^{th} observation of the j^{th} point within the network ψ_i

The estimate of variance, assuming the initial sample is selected without replacement, is given by

$$\hat{V}\left(\hat{E}(\bar{a}_{HH})\right) = \frac{(K-k)}{Kk(k-1)} \sum_{i=1}^{k} \left(u_i - \hat{E}(\bar{a}_{HH})\right)^2 \tag{2.12}$$

Example 2.1

Consider the example shown in Figure 2.4. This has been simplified to demonstrate the calculations; in reality an initial sample would typically be expected to be significantly larger, of the order of 100 or 200 points. The survey region has been overlaid with a grid of squares, with points centred within each square. Four initial points are randomly selected, without replacement, so k = 4, and the grid is 8 by 8 giving K = 64. The trigger condition is the observation of one or more animal. Three of the points meet this condition and so additional points are added; in this case the neighbourhood is defined as the points immediately above, below and to the left and right of the initial point.

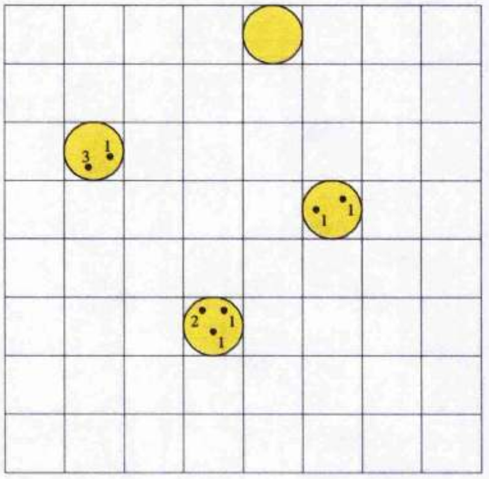


Figure 2.4: The survey region is overlaid with a grid of squares. Four initial points have been sampled, shown as solid circles. Each observation is shown as a black dot with a number beside it signifying the number of animals in the observation (the group size). For clarity, only detections are shown, although there may be other undetected animals in the surveyed area.

Figure 2.5 shows the results of adding neighbouring points to the initial points that meet the trigger condition. Three of the neighbouring points meet the trigger condition and their neighbourhoods are also added.

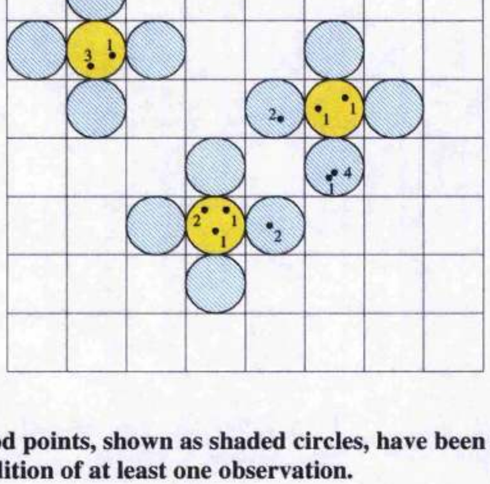


Figure 2.5: Neighbourhood points, shown as shaded circles, have been added to the initial points that meet the trigger condition of at least one observation.

Figure 2.6 shows the completed survey with all additional points added where initial points have met the trigger condition. Two of the initial points both belong to the same network. There are three distinct networks; however one of these appears twice in the analysis by the nature of the Hansen-Hurwitz estimator. Labelling the initial points 1 to 4 working from the top of the survey region to the bottom, and units within a network from 1 to n, working from left to right and top to bottom within a network, the data are summarised in Table 2.1.

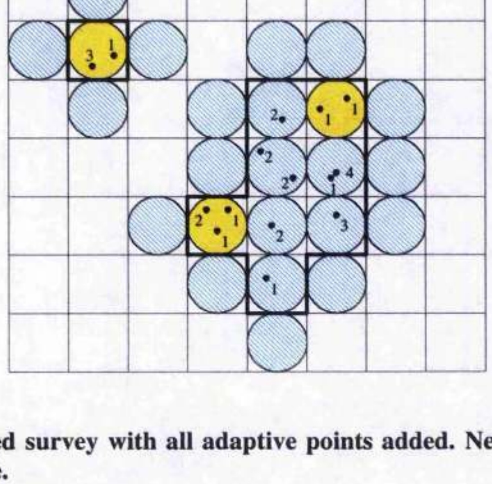


Figure 2.6: The completed survey with all adaptive points added. Networks are signified by a thick black boundary line.

Table 2.1: Summary of data for Example 2.1. Note as the 3rd and 4th initial points are both part of the same network, they share the same values for m_i, n_{ij} and a_{ijX}.

Initial Point	m_i	n_{ij}	a_{ijX}
1	$m_1=1$	$n_{11}=0$	n/a
2	$m_2=1$	$n_{21}=2$	$a_{211}=3$
			$a_{212}=1$
3	$m_3=8$	$n_{31}=1$	$a_{311}=2$
		$n_{32}=2$	$a_{321}=1$
			$a_{322}=1$
		$n_{33}=2$	$a_{331}=2$
			$a_{332}=2$
		$n_{34}=2$	$a_{341}=1$
			$a_{342}=4$
		$n_{35}=3$	$a_{351}=2$
			$a_{352}=1$
			$a_{3532}=1$
		$n_{36}=1$	$a_{361}=2$
		$n_{37}=1$	$a_{371}=3$
		$n_{38}=1$	$a_{381}=1$
4	$m_4=8$	$n_{41}=1$	$a_{411}=2$
		$n_{42}=2$	$a_{421}=1$
			$a_{422}=1$
		$n_{43}=2$	$a_{431}=2$
			$a_{432}=2$
		$n_{44}=2$	$a_{441}=1$
			$a_{442}=4$
		$n_{45}=3$	$a_{451}=2$
			$a_{452}=1$
			$a_{453}=1$
		$n_{46}=1$	$a_{461}=2$
		$n_{47}=1$	$a_{471}=3$
		$n_{48}=1$	$a_{481}=1$

Calculation of $\hat{E}(\bar{n}_{HH})$

Applying equations 2.9 and 2.10 we get

$$w_1 = \frac{1}{1}(0) = 0$$

$$w_2 = \frac{1}{1}(2) = 2$$

$$w_3 = \frac{1}{8}(1+2+2+2+3+1+1+1) = 1.625$$

$$w_4 = \frac{1}{8}(1+2+2+2+3+1+1+1) = 1.625$$

So

$$\hat{E}(\bar{n}_{HH}) = \frac{1}{4}(0+2+1.625+1.625) = 1.313$$

and

$$\hat{V}\left(\hat{E}(\bar{n}_{HH})\right)=\frac{(64-4)}{64\cdot 4\cdot(4-1)}\begin{pmatrix}(0-1.3125)^2+(2-1.3125)^2+\\ (1.625-1.3125)^2+(1.625-1.3125)^2\end{pmatrix}$$
$$=0.1868$$

Calculation of $\hat{E}(\bar{a}_{HH})$

Similarly applying equations 2.11 and 2.12 gives:

$$u_1=\frac{1}{1}(0)=0$$
$$u_2=\frac{1}{1}(1+3)=4$$
$$u_3=\frac{1}{8}((2)+(1+1)+(2+2)+(1+4)+(2+1+1)+(2)+(3)+(1))=2.875$$
$$u_4=\frac{1}{8}((2)+(1+1)+(2+2)+(1+4)+(2+1+1)+(2)+(3)+(1))=2.875$$

So

$$\hat{E}(\bar{a}_{HH})=\frac{1}{4}(0+4+2.875+2.875)=2.438$$

and

$$\hat{V}\left(\hat{E}(\bar{a}_{HH})\right)=\frac{(64-4)}{64\cdot 4\cdot(4-1)}\begin{pmatrix}(0-2.4375)^2+(4-2.4375)^2+\\ (2.875-2.4375)^2+(2.875-2.4375)^2\end{pmatrix}$$
$$=0.6848$$

RIS Design: Horvitz-Thompson-based estimators

We now use Thompson's Horvitz-Thompson-based estimator for sampling with or without replacement (equations 1.3 and 1.4), to produce unbiased estimates of expected mean number of observations per point, $\hat{E}(\bar{n}_{HT})$, and the expected mean number of animals observed per point, $\hat{E}(\bar{a}_{HT})$.

$E(\bar{n}_{HT})$ Estimate for RIS Design

Applying equations 1.3 and 1.4 to point transect surveys gives an estimate of the expected mean number of observations per point of

$$\hat{E}(\bar{n}_{HT})=\frac{1}{K}\sum_{i=1}^{\nu}\frac{\sum_{j=1}^{m_i}n_{ij}}{\alpha_i} \tag{2.13}$$

where

$$\alpha_i = 1 - \binom{K - m_i}{k} \Big/ \binom{K}{k}$$

K	is the number of points in the survey region
ν	is the number of distinct networks in the sample
m_i	is the number of points in the ith network.
n_{ij}	is the number of observations at the jth point within the ith network
α_i	is the probability that the ith network is included in the sample
k	is the number of initial points

The variance is estimated by

$$\hat{V}\left(\hat{E}(\bar{n}_{HT})\right) = \frac{1}{K^2} \sum_{i=1}^{\nu} \sum_{h=1}^{\nu} \frac{\sum_{j=1}^{m_i} n_{ij} \sum_{j=1}^{m_h} n_{hj} (\alpha_{ih} - \alpha_i \alpha_h)}{(\alpha_i \alpha_h \alpha_{ih})} \tag{2.14}$$

with

$$\alpha_{ii} = \alpha_i$$

and

$$\alpha_{ih} = 1 - \left\{ \binom{K - m_i}{k} + \binom{K - m_h}{k} - \binom{K - m_i - m_h}{k} \right\} \Big/ \binom{K}{k}$$

$E(\bar{a}_{HT})$ Estimate for RIS Design

An estimate of the expected mean number of animals observed per point is given by:

$$\hat{E}(\bar{a}_{HT}) = \frac{1}{K} \sum_{i=1}^{\nu} \frac{\sum_{j=1}^{m_i} \sum_{X=1}^{n_{ij}} a_{ijX}}{\alpha_i} \tag{2.15}$$

where

$$\alpha_i = 1 - \binom{K - m_i}{k} \Big/ \binom{K}{k}$$

a_{ijX}	is the number of animals observed in the Xth observation of the jth point in the ith network.

The variance is estimated by

$$\hat{V}\left(\hat{E}(\bar{a}_{HT})\right)=\frac{1}{K^2}\sum_{i=1}^{v}\sum_{h=1}^{v}\frac{\sum_{j=1}^{m_i}\sum_{X=1}^{n_{ij}}a_{ijX}\sum_{j=1}^{m_h}\sum_{X=1}^{n_{hj}}a_{hjX}\left(\alpha_{ih}-\alpha_i\alpha_h\right)}{\left(\alpha_i\alpha_h\alpha_{ih}\right)} \qquad (2.16)$$

with

$$\alpha_{ii}=\alpha_i$$

and

$$\alpha_{ih}=1-\left\{\binom{K-m_i}{k}+\binom{K-m_h}{k}-\binom{K-m_i-m_h}{k}\right\}\bigg/\binom{K}{k}$$

Example 2.2

Using the survey given in example 2.1, but assuming the initial points were selected without replacement, then the survey has 3 distinct networks. Labelling the networks as 1 at the top of the grid, 2 in the middle and 3 at the bottom, the data are summarised in Table 2.2. The total number of units in the survey region is K=64, the number of initial points, k=4 and the number of distinct networks, v=3.

Table 2.2: Summary of data for Example 2.2. This has three distinct networks, labelled 1 to 3 working from top to bottom of survey grid. K=64, k=4 and v=3.

Network	m_i	n_{ij}	a_{ijX}
1	$m_1=1$	$n_{11}=0$	n/a
2	$m_2=1$	$n_{21}=2$	$a_{211}=3$
			$a_{21=2}=1$
3	$m_3=8$	$n_{31}=1$	$a_{311}=2$
		$n_{32}=2$	$a_{321}=1$
			$a_{322}=1$
		$n_{33}=2$	$a_{331}=2$
			$a_{332}=2$
		$n_{34}=2$	$a_{341}=1$
			$a_{342}=4$
		$n_{35}=3$	$a_{351}=2$
			$a_{352}=1$
			$a_{3532}=1$
		$n_{36}=1$	$a_{361}=2$
		$n_{37}=1$	$a_{371}=3$
		$n_{38}=1$	$a_{381}=1$

Calculation of $\hat{E}(\bar{n}_{HT})$

From equations 2.13 and 2.14 we get:

$$\alpha_1=1-\binom{64-1}{4}\bigg/\binom{64}{4}=1-0.9375=0.0625$$

$$\alpha_2 = 1 - \binom{64-1}{4} \Big/ \binom{64}{4} = 1 - 0.9375 = 0.0625$$

$$\alpha_3 = 1 - \binom{64-8}{4} \Big/ \binom{64}{4} = 1 - 0.5781 = 0.4219$$

So

$$\hat{E}(\overline{n}_{HT}) = \frac{1}{64}\left(\frac{0}{0.0625} + \frac{2}{0.0625} + \frac{1+2+2+2+3+1+1+1}{0.4219}\right) = 0.981$$

Also

$$\alpha_{1,1} = \alpha_1$$

$$\alpha_{1,2} = \alpha_{2,1} = 1 - \left(\left(\binom{64-1}{4} + \binom{64-1}{4} - \binom{64-1-1}{4}\right)\right) \Big/ \binom{64}{4} = 0.00298$$

$$\alpha_{1,3} = \alpha_{3,1} = 1 - \left(\left(\binom{64-1}{4} + \binom{64-8}{4} - \binom{64-1-8}{4}\right)\right) \Big/ \binom{64}{4} = 0.02121$$

$$\alpha_{2,2} = \alpha_2$$

$$\alpha_{2,3} = \alpha_{3,2} = 1 - \left(\left(\binom{64-1}{4} + \binom{64-8}{4} - \binom{64-1-8}{4}\right)\right) \Big/ \binom{64}{4} = 0.02121$$

$$\alpha_{3,3} = \alpha_3$$

So

$$\hat{V}\left(\hat{E}(\overline{n}_{HT})\right) = \frac{1}{64^2}\left(\begin{array}{l} 0+0+0+ \\ 0+\left(\dfrac{2\times2\times(0.0625-0.0625\times0.0625)}{0.0625\times0.0625\times0.0625}\right)+ \\ \qquad\qquad\left(\dfrac{2\times13\times(0.02121-0.0625\times0.4219)}{0.0625\times0.4219\times0.02121}\right)+ \\ 0+\left(\dfrac{13\times2\times(0.02121-0.4219\times0.0625)}{0.4219\times0.0.0625\times0.02121}\right)+ \\ \qquad\qquad\left(\dfrac{13\times13\times(0.4219-0.4219\times0.4219)}{0.4219\times0.4219\times0.4219}\right) \end{array}\right)$$

$$= 0.251$$

Calculation of $\hat{E}(\overline{a}_{HT})$

Using equations 2.15 and 2.16 we get, with α_i and α_{ij} as in the calculation of $\hat{E}(\overline{n}_{HT})$ then

$$\hat{E}(\overline{a}_{HT}) = \frac{1}{64}\left(\frac{0}{0.0625} + \frac{3+1}{0.0625} + \frac{(2)+(1+1)+(2+2)+(1+4)+(2+1+1)+(2)+(3)+(1)}{0.4219}\right)$$
$$= 1.852$$

and

$$\hat{V}\left(\hat{E}(\overline{a}_{HT})\right) = \frac{1}{64^2}\left(\begin{array}{l} 0+0+0+ \\ 0+\left(\frac{4\times 4\times(0.0625-0.0625\times 0.0625)}{0.0625\times 0.0625\times 0.0625}\right)+ \\ \qquad\qquad\left(\frac{4\times 23\times(0.02121-0.0625\times 0.4219)}{0.0625\times 0.4219\times 0.02121}\right)+ \\ 0+\left(\frac{23\times 4\times(0.02121-0.4219\times 0.0625)}{0.4219\times 0.0.0625\times 0.02121}\right)+ \\ \qquad\qquad\left(\frac{23\times 23\times(0.4219-0.4219\times 0.4219)}{0.4219\times 0.4219\times 0.4219}\right) \end{array}\right)$$
$$= 0.9423$$

2.2.5 Adaptive Point Transect Sampling with SIS Design

We consider an initial sample in which the points have been systematically placed through the use of primary and secondary units. An example is a series of randomly selected vertical strips, the primary units, where each strip is actually a series of points, the secondary units (see Example 2.3). Thus the primary units are made up of a systematic arrangement of secondary units where the secondary units are the units of the grid. The use of the term systematic should be clarified in this context, the primary units are selected using simple random sampling and it is the arrangement of the secondary units that is systematic (Thompson and Seber, 1996: p123). However in distance sampling if the primary units (for example lines of points) have been systematically spaced on a grid, but the grid itself is randomly located on the survey region, it is common practice to treat these as a random sample.

SIS Design: Hansen-Hurwitz-based estimators

We now use Thompson's Hansen-Hurwitz-based estimators for sampling without replacement (equations 1.5 and 1.6). These are used to produce unbiased estimates of expected mean number of observations per point, $\hat{E}(\overline{n}_{SIS.HH})$, and the expected mean number of animals observed per point, $\hat{E}(\overline{a}_{SIS.HH})$.

$E(\bar{n}_{SIS.HH})$ Estimate for SIS Design

Applying equations 1.5 and 1.6 gives an estimate of the expected mean number of observations per point of

$$\hat{E}(\bar{n}_{SIS.HH}) = \frac{1}{r}\sum_{h=1}^{r} w_h \qquad (2.17)$$

where

$$w_h = \frac{1}{M}\sum_{i=1}^{\kappa_h}\frac{\sum_{j\in\psi_i} n_{ij}}{t_i}$$

r	is the number of primary units in the sample
M	is the number of secondary units (points) in each primary unit
κ_h	is the number of networks that intersect the hth primary unit
ψ_i	is the set of points in the ith network
n_{ij}	is the number of observations at the jth point of the ith network
t_i	is the number of primary units that intersect the ith network

The variance is estimated by

$$\hat{V}\left(\hat{E}(\bar{n}_{SIS.HH})\right) = \frac{s_W^2}{r}\left(1 - \frac{r}{R}\right) \qquad (2.18)$$

where

$$s_W^2 = \frac{1}{r-1}\sum_{h=1}^{r}\left(w_h - \hat{E}(\bar{n}_{SIS.HH})\right)^2$$

R	is the number of primary units in the survey region

$E(\bar{a}_{SIS.HH})$ Estimate for SIS Design

An estimate of the expected mean number of animals observed per point is given by

$$\hat{E}(\bar{a}_{SIS.HH}) = \frac{1}{r}\sum_{h=1}^{r} u_h \qquad (2.19)$$

where

$$u_h = \frac{1}{M}\sum_{i=1}^{\kappa_h}\frac{\sum_{j\in\psi_i}\sum_{X=1}^{n_{ij}} a_{ijX}}{t_i}$$

a_{ijX}	is the number of animals in the Xth observation of the jth point of the ith network

The variance is estimated by

$$\hat{V}\left(\hat{E}(\overline{a}_{SIS.HH})\right) = \frac{s_W^2}{r}\left(1 - \frac{r}{R}\right) \tag{2.20}$$

where

$$s_W^2 = \frac{1}{r-1}\sum_{h=1}^{r}\left(u_h - \hat{E}(\overline{a}_{SIS.HH})\right)^2$$

Example 2.3

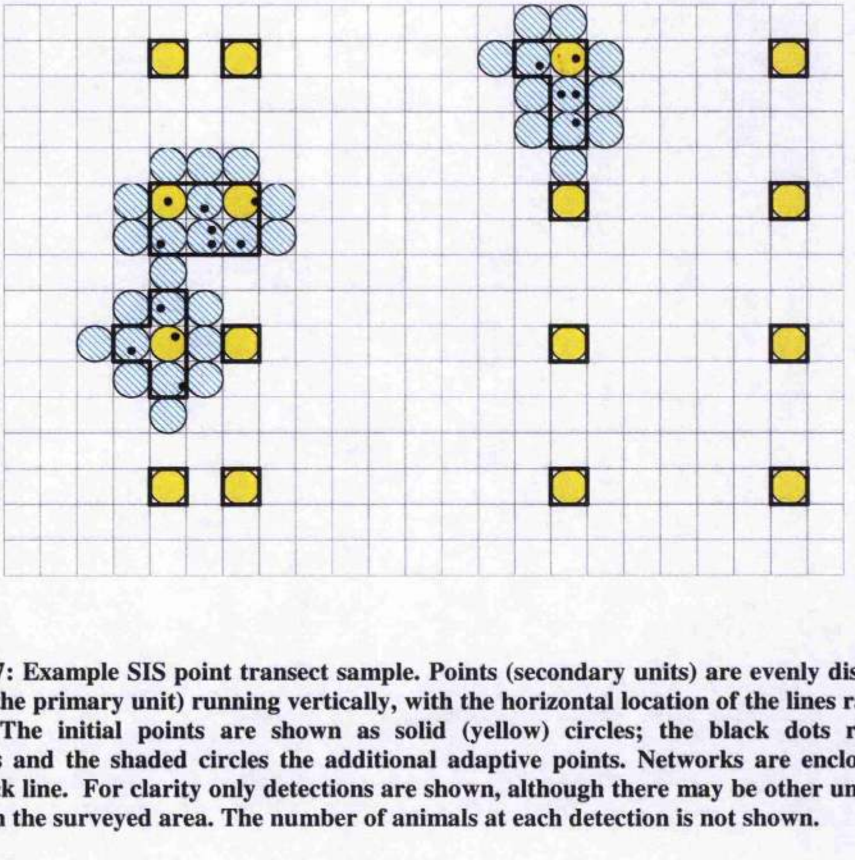


Figure 2.7: Example SIS point transect sample. Points (secondary units) are evenly distributed on lines (the primary unit) running vertically, with the horizontal location of the lines randomly selected. The initial points are shown as solid (yellow) circles; the black dots represent detections and the shaded circles the additional adaptive points. Networks are enclosed in a thick black line. For clarity only detections are shown, although there may be other undetected animals in the surveyed area. The number of animals at each detection is not shown.

The example shown in Figure 2.7 is intended only to demonstrate the approach and does not represent a realistic survey. Here we have a 23 by 16 units survey grid, with systematically distributed secondary units and four randomly located primary units (r=4), selected without replacement. The potential points are centred within each square unit. Each primary unit is a vertical line consisting of four evenly spaced secondary units (M=4), the points. The primary units are randomly spaced horizontally and share a common random vertical start position. The grid is 16 units high and the secondary units are spaced at 4 unit intervals, thus with 4 secondary units in a primary unit, there are 4 possible vertical start positions. The grid is 23 units wide and so there are 4x23 potential primary units (R=92) within the survey region.

The initial points are shown as solid circles and the adaptive points as shaded. For clarity, only detections are shown, as solid black dots, although there may be other, undetected, animals in the surveyed area. Each detection relates to one or more animals with the group sizes given in Table 2.3.

Table 2.3: Summary of data for Example 2.3. r=4, M=4 and ν=3.

Network	t_i	n_{ij}	a_{ijX}
1	$t_1=1$	$n_{1,1}=0$	n/a
2	$t_2=1$	$n_{2,1}=0$	n/a
3	$t_3=4$	$n_{3,1}=1$	$a_{3,1,1}=1$
		$n_{3,2}=1$	$a_{3,2,1}=3$
		$n_{3,3}=2$	$a_{3,3,1}=3$
			$a_{3,3,2}=3$
		$n_{3,4}=1$	$a_{3,4,1}=2$
4	$t_4=1$	$n_{4,1}=0$	n/a
5	$t_5=6$	$n_{5,1}=1$	$a_{5,1,1}=3$
		$n_{5,2}=1$	$a_{5,2,1}=1$
		$n_{5,3}=1$	$a_{5,3,1}=3$
		$n_{5,4}=1$	$a_{5,4,1}=2$
		$n_{5,5}=2$	$a_{5,5,1}=3$
			$a_{5,5,2}=2$
		$n_{5,6}=1$	$a_{5,6,1}=3$
6	$t_6=1$	$n_{6,1}=0$	n/a
7	$t_7=1$	$n_{7,1}=0$	n/a
8	$t_8=4$	$n_{8,1}=1$	$a_{8,1,1}=1$
		$n_{8,2}=1$	$a_{8,2,1}=1$
		$n_{8,3}=1$	$a_{8,3,1}=2$
		$n_{8,4}=1$	$a_{8,4,1}=3$
9	$t_9=1$	$n_{9,1}=0$	n/a
10	$t_{10}=1$	$n_{10,1}=0$	n/a
11	$t_{11}=1$	$n_{11,1}=0$	n/a
12	$t_{12}=1$	$n_{12,1}=0$	n/a
13	$t_{13}=1$	$n_{13,1}=0$	n/a
14	$t_{14}=1$	$n_{14,1}=0$	n/a
15	$t_{15}=1$	$n_{15,1}=0$	n/a

There are 15 distinct networks (ν=15) in the sample. Only three of these include any detections; the rest all consist of a single initial point. The networks are labelled 1 to 15 working from left to right and top to bottom. Thus for the networks with detections, network 3 is in the upper right of the survey region, network 5 in the middle on the left and network 8 below it. Units within each network are numbered working from left to right and top to bottom within each network. Four primary units are sampled, so r=4, and there are four secondary units in each primary unit, so M=4. Labelling the primary units, 1 to 4 from left to right, then the number of networks (with a non-zero count of animals) that intersect each primary unit are: $\kappa_1=2$, $\kappa_2=1$, $\kappa_3=1$ and $\kappa_4=0$.

Of the 92 potential locations for the primary units, 4 of these intersect network 3 and so $t_3=4$. Similarly for networks 5 and 8 we have $t_5=6$ and $t_8=4$. For each of the networks with no detections, as they are a single point in size, there is only one primary unit that intersects each.

Calculation of $\hat{E}(\bar{n}_{SIS.HH})$

Using equations 2.17 and 2.18 we get

$$w_1 = \frac{1}{4}\left(\frac{0}{1} + \frac{1+1+1+1+2+1}{6} + \frac{1+1+1+1}{4} + \frac{0}{1}\right) = 0.542$$

$$w_2 = \frac{1}{4}\left(\frac{0}{1} + \frac{1+1+1+1+2+1}{6} + \frac{0}{1} + \frac{0}{1}\right) = 0.292$$

$$w_3 = \frac{1}{4}\left(\frac{1+1+2+1}{4} + \frac{0}{1} + \frac{0}{1} + \frac{0}{1}\right) = 0.313$$

$$w_4 = \frac{1}{4}\left(\frac{0}{1} + \frac{0}{1} + \frac{0}{1} + \frac{0}{1}\right) = 0.0$$

So that

$$\hat{E}(\bar{n}_{SIS.HH}) = \frac{1}{4}(0.542 + 0.292 + 0.313 + 0) = 0.286$$

and

$$\begin{aligned} s_w^2 &= \frac{1}{4-1}\left((0.542-0.286)^2 + (0.292-0.286)^2 + (0.313-0.286)^2 + (0-0.286)^2\right) \\ &= 0.0493 \end{aligned}$$

So

$$\hat{V}\left(\hat{E}(\bar{n}_{SIS.HH})\right) = \frac{0.0.0493}{4}\left(1 - \frac{4}{92}\right) = 0.0118$$

Calculation of $\hat{E}(\bar{a}_{SIS.HH})$

From equations 2.19 and 2.20 we have

$$u_1 = \frac{1}{4}\left(\frac{0}{1} + \frac{3+1+3+2+(3+2)+3}{6} + \frac{1+1+2+3}{4} + \frac{0}{1}\right) = 1.146$$

$$u_2 = \frac{1}{4}\left(\frac{0}{1} + \frac{3+1+3+2+(3+2)+3}{6} + \frac{0}{1} + \frac{0}{1}\right) = 0.708$$

$$u_3 = \frac{1}{4}\left(\frac{1+3+(3+3)+2}{4} + \frac{0}{1} + \frac{0}{1} + \frac{0}{1}\right) = 0.750$$

$$u_4 = \frac{1}{4}\left(\frac{0}{1} + \frac{0}{1} + \frac{0}{1} + \frac{0}{1}\right) = 0.0$$

So

$$\hat{E}(\overline{a}_{SIS.HH}) = \frac{1}{4}(1.146 + 0.708 + 0.750 + 0.0) = 0.651$$

and

$$\begin{aligned} s_w^2 &= \frac{1}{4-1}\left((1.146-0.651)^2 + (0.708-0.651)^2 + (0.750-0.651)^2 + (0-0.651)^2\right) \\ &= 0.2273 \end{aligned}$$

So

$$\hat{V}\left(\hat{E}(\overline{a}_{SIS.HH})\right) = \frac{0.2273}{4}\left(1 - \frac{4}{92}\right) = 0.0543$$

SIS Design: Horvitz-Thompson-based estimator

The fourth set of estimators use Thompson's Horvitz-Thompson-based estimator (equations 1.7 and 1.8), with the primary units selected without replacement.

$E(\overline{n}_{SIS.HT})$ Estimate for SIS Design

From equations 1.7 and 1.8, an estimate of the expected mean number of observations per point is

$$\hat{E}(\overline{n}_{SIS.HT}) = \frac{1}{MR}\sum_{i=1}^{v}\frac{\sum_{j=1}^{m_i} n_{ij}}{\alpha_i} \tag{2.21}$$

where

$$\alpha_i = 1 - \binom{R-t_i}{r} \Big/ \binom{R}{r}$$

M	is the number of secondary units (points) in each primary unit
R	is the number of primary units in the survey region
v	is the number of distinct networks in the sample

m_i	is the number of points in the i[th] network
n_{ij}	is the number of observations at the j[th] point of the i[th] network
α_i	is the probability that the i[th] network is included in the sample
t_i	is the number of primary units that intersect the i[th] network
r	is the number of primary units in the sample

with variance

$$\hat{V}\left(\hat{E}(\overline{n}_{SIS.HT})\right)=\frac{1}{M^2R^2}\sum_{i=1}^{v}\sum_{h=1}^{v}\frac{\sum_{j=1}^{m_i}\sum_{j'=1}^{m_h}n_{ij}n_{ij'}(\alpha_{ih}-\alpha_i\alpha_h)}{(\alpha_i\alpha_h\alpha_{ih})} \tag{2.22}$$

where

$$\alpha_{ih}=1-\left\{\binom{R-t_i}{r}+\binom{R-t_h}{r}-\binom{R-t_i-t_h+t_{ih}}{r}\right\}\Big/\binom{R}{r}$$

$$\alpha_{ii}=\alpha_i$$

and

t_{ih}	is the number of primary units that intersect networks i and h.

$E(\overline{a}_{SIS.HT})$ **Estimate for SIS Design**

An estimate of the mean total number of animals observed per point is

$$\hat{E}(\overline{a}_{SIS.HT})=\frac{1}{MR}\sum_{i=1}^{v}\frac{\sum_{j=1}^{m_i}\sum_{X=1}^{n_{ij}}a_{ijX}}{\alpha_i} \tag{2.23}$$

where

$$\alpha_i=1-\binom{R-t_i}{r}\Big/\binom{R}{r}$$

a_{ijX}	is the number of animals in the X[th] observation of the j[th] point of the i[th] network

with variance

$$\hat{V}\left(\hat{E}(\overline{a}_{SIS.HT})\right)=\frac{1}{M^2R^2}\sum_{i=1}^{v}\sum_{h=1}^{v}\frac{\sum_{j=1}^{m_i}\sum_{X=1}^{n_{ij}}\sum_{j'=1}^{m_h}\sum_{X'=1}^{n_{ij'}}a_{ijX}\,a_{hj'X'}(\alpha_{ih}-\alpha_i\alpha_h)}{(\alpha_i\alpha_h\alpha_{ih})} \tag{2.24}$$

where

$$\alpha_{ih} = 1 - \left\{ \binom{R - t_i}{r} + \binom{R - t_h}{r} - \binom{R - t_i - t_h + t_{ih}}{r} \right\} \Big/ \binom{R}{r}$$

$$\alpha_{ii} = \alpha_i$$

Example 2.4

We again use the survey given in Figure 2.7, this time assuming the horizontal location of the primary units was selected without replacement. The summary of the data given in Table 2.3 is also re-used for this example.

Calculation of $\hat{E}(\bar{n}_{SIS.HT})$

We now apply equations 2.21 and 2.22. As the networks with no observations evaluate to zero and so make no contribution to the estimates, they have been omitted from the calculations to simplify presentation. There are only two networks in the sample, where a primary unit can intersect both networks, networks 5 and 8. For the 5th column (from the left of the grid) the primary units starting in the 2nd and 3rd rows (down from the top) both intersect networks 5 and 8, so that $t_{5,8}=2=t_{8,5}$. For all other network combinations $t_{ih}=0$.

$$\alpha_1 = 1 - \binom{92-1}{4} \Big/ \binom{92}{4} = 0.043$$

$$\alpha_2 = \alpha_4 = \alpha_6 = \alpha_7 = \alpha_9 = \alpha_{10} = \alpha_{11} = \alpha_{12} = \alpha_{13} = \alpha_{14} = \alpha_{15} = \alpha_1 = 0.043$$

$$\alpha_3 = 1 - \binom{92-4}{4} \Big/ \binom{92}{4} = 0.175$$

$$\alpha_5 = 1 - \binom{92-6}{4} \Big/ \binom{92}{4} = 0.240$$

$$\alpha_8 = 1 - \binom{92-4}{4} \Big/ \binom{92}{4} = 0.175$$

So

$$\hat{E}(\bar{n}_{SIS.HT}) = \frac{1}{4 \times 92} \left(\frac{(1+1+2+1)}{0.175} + \frac{(1+1+1+1+2+1)}{0.240} + \frac{(1+1+1+1)}{0.175} \right) = 0.219$$

and

$$\alpha_{3,3} = \alpha_3$$

$$\alpha_{3,5} = \alpha_{5,3} = 1 - \left(\binom{92-4}{4} + \binom{92-6}{4} - \binom{92-4-6+2}{4} \right) \Big/ \binom{92}{4} = 0.0960$$

$$\alpha_{3,8} = \alpha_{8,3} = 1 - \left(\binom{92-4}{4} + \binom{92-4}{4} - \binom{92-4-4+0}{4} \right) \Big/ \binom{92}{4} = 0.0214$$

$$\alpha_{5,5} = \alpha_5$$

$$\alpha_{5,8} = \alpha_{8,5} = 1 - \left(\binom{92-6}{4} + \binom{92-4}{4} - \binom{92-6-4+0}{4} \right) \Big/ \binom{92}{4} = 0.0314$$

$$\alpha_{8,8} = \alpha_8$$

So

$$\hat{V}\left(\hat{E}(\overline{n}_{SIS.HT})\right) = \frac{1}{4^2 \times 92^2} \left(\begin{array}{c} \frac{5\times5\times(0.175-0.175\times0.175)}{0.175\times0.175\times0.175} + \frac{5\times7\times(0.0960-0.175\times0.240)}{0.175\times0.240\times0.0960} + \\ \frac{5\times4\times(0.0214-0.175\times0.175)}{0.175\times0.175\times0.0214} + \\ \frac{7\times5\times(0.0960-0.240\times0.175)}{0.240\times0.175\times0.0960} + \frac{7\times7\times(0.240-0.240\times0.240)}{0.240\times0.240\times0.240} + \\ \frac{7\times4\times(0.0314-0.240\times0.175)}{0.240\times0.175\times0.0314} + \\ \frac{4\times5\times(0.0214-0.175\times0.175)}{0.175\times0.175\times0.0214} + \frac{4\times7\times(0.0314-0.175\times0.240)}{0.175\times0.240\times0.0314} + \\ \frac{4\times4\times(0.175-0.175\times0.175)}{0.175\times0.175\times0.175} \end{array} \right)$$

$$= 0.0124$$

Calculation of $\hat{E}(\overline{a}_{SIS.HT})$

Applying equations 2.23 and 2.24, with α_l and α_{lh} as in the calculation of $E(\overline{n}_{SIS.HT})$ and as before ignoring networks with no detections, we get

$$\hat{E}(\overline{a}_{SIS.HT}) = \frac{1}{4\times92}\left(\frac{(1+3+3+3+2)}{0.175} + \frac{(3+1+3+2+3+2+3)}{0.240} + \frac{(1+1+2+3)}{0.175}\right)$$
$$= 0.488$$

and

$$\hat{V}\left(\hat{E}(\overline{a}_{SIS,HT})\right)=\frac{1}{4^2\times 92^2}\left(\begin{array}{l}\frac{12\times12\times(0.175-0.175\times0.175)}{0.175\times0.175\times0.175}+\frac{12\times17\times(0.0960-0.175\times0.240)}{0.175\times0.240\times0.0960}+\\ \frac{12\times7\times(0.0214-0.175\times0.175)}{0.175\times0.175\times0.0214}+\\ \frac{17\times12\times(0.0960-0.240\times0.175)}{0.240\times0.175\times0.0960}+\frac{17\times17\times(0.240-0.240\times0.240)}{0.240\times0.240\times0.240}+\\ \frac{17\times7\times(0.0314-0.240\times0.175)}{0.240\times0.175\times0.0314}+\\ \frac{7\times12\times(0.0214-0.175\times0.175)}{0.175\times0.175\times0.0214}+\frac{7\times17\times(0.0314-0.175\times0.240)}{0.175\times0.240\times0.0314}+\\ \frac{7\times7\times(0.175-0.175\times0.175)}{0.175\times0.175\times0.175}\end{array}\right)$$

$$=0.0753$$

2.2.6 Unequal Probability of Detection

For many practical applications the probability of detection can be considered equal for all animals in the survey. By this we mean the detection function is the same, although probability of detection is likely to vary with distance. The basic approach used so far in this chapter, assumes the probability of detection of an object is purely dependent on its distance from the point and so observations are pooled across all points, initial and adaptive, to produce an estimate of $h(0)$. However there are likely to be many other factors, such as group size (large groups are more likely to be seen at greater distances), weather, habitat, altitude, observer experience, observer awareness, etc. Some of these can be addressed by stratifying, for example by weather conditions, altitude or habitat, however in recent years much work has been done to allow the inclusion of covariates in estimating the detection function. Developments in distance sampling (Borchers 1996; Borchers *et al.*, 1998a; Buckland *et al.*, 2002; Marques 2001; Strindberg 2001) are now encouraging the use of these as covariates in the $h(0)$ estimation.

Here we extend the earlier derived Horvitz-Thompson-based estimators (equations 2.13 – 2.16) to allow for each observation to have its own probability of detection.

RIS Design: Horvitz-Thompson-based Estimate for Unequal Probability of Detection

Thompson and Seber (1996: p228) and Thompson and Seber (1994), present formulae to deal with an adaptive cluster sampling design, where each object has a unique probability of detection. For a random initial sample and utilising a Horvitz-

Thompson-based estimator, an estimate of the total population for the survey region, $\tilde{\tau}$, is given by:

$$\tilde{\tau} = \sum_{i=1}^{\nu} \frac{\hat{u}_i^*}{\alpha_i} \tag{2.25}$$

where

$$\hat{u}_i^* = \sum_{X=1}^{m_i} \frac{y_{iX}}{\hat{g}_{iX}}$$

ν	is the number of distinct networks in the sample
α_i	is the inclusion probability
m_i	is the number of observations in the ith network
y_{iX}	is the variable of interest for the Xth observation in the ith network
$\hat{g}_{iX}$	is an estimate of the probability of detection for the Xth observation of the ith network

Thompson and Seber estimate the variance as consisting of 3 main components. Simplistically this can be considered the variance from the adaptive estimate of n, the variance from the probability of detection and finally the variance due to estimation of the probability of detection. Thus

$$\hat{\text{var}}[\tilde{\tau}] = \hat{\upsilon}_1 + \hat{\upsilon}_2 + \hat{\upsilon}_3 \tag{2.26}$$

where

$$\hat{\upsilon}_1 = \sum_{i=1}^{\nu} \sum_{i'=1}^{\nu} \hat{u}_i^* \hat{u}_{i'}^* \left(\frac{\alpha_{ii'} - \alpha_i \alpha_{i'}}{\alpha_{ii'} \alpha_i \alpha_{i'}} \right)$$

$$\hat{\upsilon}_2 = \sum_{i=1}^{\nu} \frac{1}{\alpha_i} \sum_{X=1}^{m_i} \frac{1 - \hat{g}_{iX}}{\hat{g}_{iX}^2} y_{iX}^2$$

$$\hat{\upsilon}_3 = \sum_{i=1}^{\nu} \sum_{X=1}^{m_i} \sum_{i'=1}^{\nu} \sum_{X'=1}^{m_{i'}} \frac{y_{iX} \, y_{i'X'}}{\alpha_{ii'} \hat{g}_{iX}^2 \hat{g}_{i'X'}^2} \hat{\text{cov}}(\hat{g}_{iX}, \hat{g}_{i'X'})$$

Covariates and *h*(0).

Following the work of Borchers (1996), Borchers *et al.* (1998a), and Marques (2001) let $\mathbf{z}_X$ ($\mathbf{z}_X = z_{X1}, z_{X2}, \ldots, z_{Xq}$) be a matrix of the associated covariates for the Xth observation so that its detection function is $g(r \mid \mathbf{z}_X)$, at distance r from the point.

We assume that for each point the objects are uniformly distributed, with respect to distance from the point, within the sampling circle of radius R; the truncation distance. This is a stronger assumption than is necessary. For a line transect survey, Fewster and Buckland (in preparation) demonstrate that it is only required for the object distribution, with respect to the trackline, to have a linear probability distribution function (pdf). Based on this it can be shown that the point transect requirement is for the object distribution over the sampling circle to be such that for any cross section through the point, the object distances have a linear probability density function (pdf). The intuitive explanation is that any overabundance of objects in one part of the circle is compensated for by an under abundance in the rest of the circle.

Thus with truncation distance R, the probability that an animal is at distance r or less from the point is

$$P(R \leq r) = \frac{Area\ of\ the\ circle\ at\ radius\ r}{Area\ of\ the\ circle\ at\ radius\ R} = \frac{\pi r^2}{\pi R^2} = \frac{r^2}{R^2}, \qquad 0 \leq r \leq R$$

So the probability density function of the distance r to each animal from the point is

$$u(r) = \frac{d\{P(R \leq r)\}}{dr} dr = \frac{2r}{R^2} \tag{2.27}$$

The probability p_X, that object X is detected within a circular plot of truncation radius R, and conditional on the observed values $\mathbf{z}_X$, is given by

$$\begin{aligned} p_X &= E_r[g(r \mid \mathbf{z}_X)] \\ &= \int_0^R g(r \mid \mathbf{z}_X) \cdot u(r \mid \mathbf{z}_X)\, dr \end{aligned}$$

where

$g(r \mid \mathbf{z}_X)$ is the multivariate detection function for the Xth observation at distance r with covariates $\mathbf{z}_X$

$u(r \mid \mathbf{z}_X)$ is the joint density function for distance r and covariates $\mathbf{z}_X$ for the Xth observation

Assuming the r and $\mathbf{z}_X$ are independent so that

$$p_X = \int_0^R g(r \mid \mathbf{z}_X)\cdot u(r)\, dr$$

and then substituting equation 2.27 for $u(r)$ we get

$$\begin{aligned} p_X &= \int_0^R g(r \mid \mathbf{z}_X)\cdot \frac{2r}{R^2}\, dr \\ &= \frac{2}{R^2}\int_0^R g(r \mid \mathbf{z}_X)\cdot r\, dr \end{aligned} \tag{2.28}$$

We now define the probability density function $f(r \mid \mathbf{z}_X)$ in terms of the detection function $g(r \mid \mathbf{z}_X)$ by considering a ring, of incremental width dr at distance r from the point. Thus

$$\begin{aligned} f(r \mid \mathbf{z}_X)dr &= \mathrm{pr}[\{\text{object in } (r, r+dr) \mid \mathbf{z}_X\} \mid \{\text{object detected} \mid \mathbf{z}_X\}] \\ &= \frac{\mathrm{pr}[\{\text{object in } (r, r+dr) \mid \mathbf{z}_X\} \cap \{\text{object detected} \mid \mathbf{z}_X\}]}{\mathrm{pr}\{\text{object detected} \mid \mathbf{z}_X\}} \\ &= \frac{\mathrm{pr}[\{\text{object detected} \mid \mathbf{z}_X\} \mid \{\text{object in } (r, r+dr) \mid \mathbf{z}_X\}]\cdot \mathrm{pr}[\{\text{object in } (r, r+dr) \mid \mathbf{z}_X\}]}{p_X} \\ &= \frac{\mathrm{pr}[\{\text{object detected} \mid \mathbf{z}_X\} \mid \{\text{object in } (r, r+dr) \mid \mathbf{z}_X\}]\cdot \mathrm{pr}[\{\text{object in } (r, r+dr)\}]}{p_X} \\ &= \frac{g(r \mid \mathbf{z}_X)\cdot \left(\dfrac{2\pi r\, dr}{\pi R^2}\right)}{\dfrac{2}{R^2}\displaystyle\int_0^R g(r \mid \mathbf{z}_X)\cdot r\, dr} \\ &= \frac{g(r \mid \mathbf{z}_X)\cdot r\, dr}{\displaystyle\int_0^R g(r \mid \mathbf{z}_X)\cdot r\, dr} \end{aligned}$$

so that

$$f(r \mid \mathbf{z}_X) = \frac{r\cdot g(r \mid \mathbf{z}_X)}{\displaystyle\int_0^R g(r \mid \mathbf{z}_X)\cdot r\, dr}$$

Thus

$$\frac{f(r \mid \mathbf{z}_X)}{r} = \frac{g(r \mid \mathbf{z}_X)}{\displaystyle\int_0^R g(r \mid \mathbf{z}_X)\cdot r\, dr}$$

so that if detection on the point is certain, $g(0 \mid \mathbf{z}_X) = 1$, then the slope of the derivative of the probability density function of detection distances evaluated at zero is

$$h(0 \mid \mathbf{z}_X) = \lim_{r \to 0} \frac{f(r \mid \mathbf{z}_X)}{r} = \frac{1}{\int_0^R g(r \mid \mathbf{z}_X) \cdot r \, dr}$$

and so an estimate of $h(0 \mid \mathbf{z}_X)$ is given by

$$\hat{h}(0 \mid \mathbf{z}_X) = \frac{1}{\int_0^R \hat{g}(r \mid \mathbf{z}_X) \cdot r \, dr}$$

Substituting this in equation 2.28, if detection on the point is certain, an estimate of p_X is given by

$$\hat{p}_X = \frac{2}{R^2 \cdot \hat{h}(0 \mid \mathbf{z}_X)} \tag{2.29}$$

where

$\hat{h}(0 \mid \mathbf{z}_X)$ is an estimate of the derivative of the probability density function of detection distances, given $\mathbf{z}_X$, evaluated at zero

Adaptive Point Transect Sampling Estimator with Unequal Probability of Detection

We now merge Thompson's adaptive formula (equations 2.25 and 2.26) with the estimate of p_X based on $h(0)$ using covariates (equation 2.29) to create an adaptive point transect estimate where there is unequal probability of detection.

We estimate the number of animal groups in the surveyed area, τ. This new estimate differs from the earlier estimates in two ways. First it includes the probability of detection, and so is an estimate of the number of animal groups, rather than the number of animal groups detected. Second, we are obtaining an estimate of the number of animal groups across all K points in the survey region, the surveyed area. The surveyed area is a notional area and may be larger or smaller than the actual survey region. If, for example, plots are larger than the grid units and so overlap (Figure 2.3), then any animal groups in overlapping plots will be counted once for each plot they fall within. Thus, τ may be larger or smaller than N_g the number of animals in the survey region and has to be scaled by the ratio of the two areas, the surveyed area and the survey region area (see equation 2.3).

With *h(0)* estimated using covariates, then in an adaptive distance sampling context, an estimate of the number of animal groups in the surveyed area is given by

$$\hat{\tau}_{UP.HT} = \sum_{i=1}^{v} \frac{\hat{u}_i^*}{\alpha_i} \tag{2.30}$$

where

$$\hat{u}_i^* = \sum_{j=1}^{m_i} \sum_{X=1}^{n_{ij}} \frac{1}{\hat{p}_{ijX}}$$

$$\alpha_i = 1 - \binom{K - m_i}{k} \Big/ \binom{K}{k}$$

K	is the number of points in the survey region
v	is the number of distinct networks in the sample
α_i	is the inclusion probability of the i^{th} network
m_i	is the number of points in the i^{th} network
n_{ij}	is the number of observations at the j^{th} point of the i^{th} network
$\hat{p}_{ijX}$	is an estimate of the probability of detection of the X^{th} observation of the j^{th} point of the i^{th} network
k	is the number of initial survey points

The accompanying variance estimate is given by

$$\begin{aligned}\hat{V}(\hat{\tau}_{UP.HT}) = &\sum_{i=1}^{v} \sum_{i'=1}^{v} \hat{u}_i^* \hat{u}_{i'}^* \left(\frac{\alpha_{ii'} - \alpha_i \alpha_{i'}}{\alpha_{ii'} \alpha_i \alpha_{i'}} \right) \\ &+ \sum_{i=1}^{v} \frac{1}{\alpha_i} \sum_{j=1}^{m_i} \sum_{X=1}^{n_{ij}} \frac{1 - \hat{p}_{ijX}}{\hat{p}_{ijx}^2} \\ &+ \sum_{i=1}^{v} \sum_{j=1}^{m_i} \sum_{X}^{n_{ij}} \sum_{i'=1}^{v} \sum_{j'=1}^{m_{i'}} \sum_{X'}^{n_{i'j'}} \frac{1}{\alpha_{ii'} \hat{p}_{ijx}^2 \hat{p}_{i'j'x'}^2} \widehat{\mathrm{cov}}\left(\hat{p}_{ijX}, \hat{p}_{i'j'X'}\right)\end{aligned} \tag{2.31}$$

where

$$\alpha_{ii} = \alpha_i$$

and

$$\alpha_{ih} = 1 - \left\{ \binom{K - m_i}{k} + \binom{K - m_h}{k} - \binom{K - m_i - m_h}{k} \right\} \Big/ \binom{K}{k}$$

Substituting equation 2.28 for the $\hat{p}_{ijX}$ in equation 2.30 gives

$$\hat{\tau}_{UP.HT} = \frac{R^2}{2}\sum_{i=1}^{v}\frac{\hat{u}_i^*}{\alpha_i} \tag{2.32}$$

where

$$\hat{u}_i^* = \sum_{j=1}^{m_i}\sum_{X=1}^{n_{ij}}\left(\frac{1}{\int_0^R \hat{g}(r \mid \mathbf{z}_{ijX})\cdot r\, dr}\right)$$

R is the truncation radius

If probability of detection on the point is certain, $g(0 \mid \mathbf{z}_{ijX}) = 1$, then from equation 2.29 this can be further simplified to

$$\hat{\tau}_{UP.HT} = \frac{R^2}{2}\sum_{i=1}^{v}\left(\sum_{j=1}^{m_i}\sum_{X=1}^{n_{ij}}\hat{h}(0 \mid \mathbf{z}_{ijX})\Big/\alpha_i\right) \tag{2.33}$$

where

$\hat{h}(0 \mid \mathbf{z}_{ijX})$ is an estimate of the derivative of the probability density function of detection distances, for the X^{th} observation of the j^{th} point of the i^{th} network, evaluated at zero

Similarly an estimate of the number of animals in the K circular plots of radius R, is given by

$$\hat{\delta}_{UP.HT} = \frac{R^2}{2}\sum_{i=1}^{v}\frac{\hat{w}_i^*}{\alpha_i} \tag{2.34}$$

where

$$\hat{w}_i^* = \sum_{j=1}^{m_i}\sum_{X=1}^{n_{ij}}\left(\frac{s_{ijX}}{\int_0^R g(r \mid \mathbf{z}_{ijX})\cdot r\, dr}\right)$$

s_{ijX} is the group size for the X^{th} observation of the j^{th} point of the i^{th} network

with variance estimated by

$$\hat{V}\left(\hat{\delta}_{UP.HT}\right)=\sum_{i=1}^{v}\sum_{i'=1}^{v}\hat{w}_i^*\hat{w}_{i'}^*\left(\frac{\alpha_{ii'}-\alpha_i\alpha_{i'}}{\alpha_{ii'}\alpha_i\alpha_{i'}}\right)$$
$$+\sum_{i=1}^{v}\frac{1}{\alpha_i}\sum_{j=1}^{m_i}\sum_{X=1}^{n_{ij}}\frac{1-\hat{p}_{ijX}}{\hat{p}_{ijx}^2}s_{ijX} \tag{2.35}$$
$$+\sum_{i=1}^{v}\sum_{j=1}^{m_i}\sum_{X}^{n_{ij}}\sum_{i'=1}^{v}\sum_{j'=1}^{m_{i'}}\sum_{X'}^{n_{i'j'}}\frac{s_{ijX}\,s_{i'j'X'}}{\alpha_{ii'}\hat{p}_{ijx}^2\,\hat{p}_{i'j'x'}^2}\mathrm{c\hat{o}v}\left(\hat{p}_{ijX}\,,\hat{p}_{i'j'X'}\right)$$

where, as before

$$\hat{p}_{ijX}=\frac{2}{R^2}\int_0^R g\left(r\,|\,\mathbf{z}_{ijX}\right)\cdot r\,dr$$

As with the $\hat{\tau}_{UP.HT}$ estimate, this can be further simplified if $g\left(0\,|\,\mathbf{z}_{ijX}\right)=1$, by replacing the estimate of p_{ijX} by its simplified form in equation 2.29.

Density Estimators with Unequal Probability of Detection

$\hat{\tau}_{UP.HT}$ is an estimate of the number of animal groups in the surveyed area, accounting for the probability of detection, so that the group density is simply obtained by dividing by the surveyed area. Now the surveyed area is given by the total number of points in the survey region multiplied by the area of each plot, which is a circle of radius R; so that the surveyed area is $\pi R^2 K$. It follows that an estimate of the density of animal groups is given by

$$\hat{D}_{UP.g}=\frac{\hat{\tau}_{UP.HT}}{\pi R^2 K} \tag{2.36}$$

where

R	is the truncation distance
K	the number of points in the survey region

with variance estimated by

$$\hat{V}\left(\hat{D}_{UP.g}\right)=\frac{\hat{V}\left(\hat{\tau}_{UP.HT}\right)}{\pi^2 R^4 K^2} \tag{2.37}$$

Similarly the individual animal density estimate is

$$\hat{D}_{UP}=\frac{\hat{\delta}_{UP.HT}}{\pi R^2 K} \tag{2.38}$$

and an estimate of variance is given by

$$\hat{V}\left(\hat{D}_{UP}\right)=\frac{\hat{V}\left(\hat{\delta}_{UP.HT}\right)}{\pi^2 R^4 K^2} \tag{2.39}$$

For many surveys the observed group size is also likely to be an estimate, although at this point no attempt has been made to include this factor into what is already a complex estimator. The variance estimate above makes many assumptions about the independence of variables and is likely to underestimate the variance, thus in practice a bootstrap based estimate will be preferred. This is discussed later in the chapter.

2.3 Grid Design

Thompson typically uses a grid of squares with the neighbourhood being the adjacent squares to the north, south, east and west of the sampled unit. However any pattern is appropriate as long as there is symmetry so that the inclusion of any unit in a network will also include all other units of the network. Many designs are possible with point transect surveys, where the plot around the point can be fitted either within or outside the chosen unit shape. For example Figure 2.8 shows point transect sampling plots overlaid on triangular, square and hexagonal units, all three of which can easily be built into a grid design.

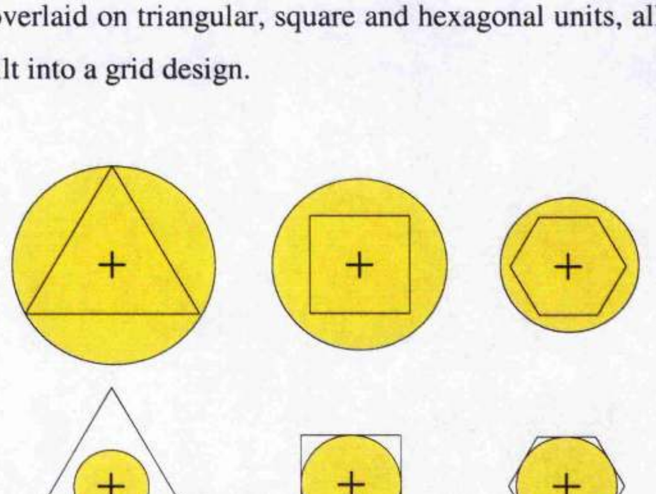


Figure 2.8: Point transect plots truncated both inside and outside of a variety of sample unit shapes. The sampling area is signified by the shading.

Many factors need to be taken into account when selecting the unit shape. The coverage can be maximised by use of a shape closely matched to a circle, such as a hexagon. The use of a large neighbourhood will increase the number of edge units, which will potentially reduce the gain from the adaptive approach, as the edge units make a limited contribution in comparison to units within a network. Not all adjoining units have to be part of the neighbourhood, and not all units in a neighbourhood have to be adjoining, the only requirement is the symmetry of inclusion.

If detections are made independently then the sampling circles can overlap. However it may be difficult to ensure independence and thus no overlap will typically be the preferred approach. For sampling circles that touch or are in close proximity survey procedures will need to be in place to ensure objects do not get counted at both points when they only reside within the truncation radius of one point. As with conventional point transect sampling, if not removed, this type of double counting will bias results, causing an overestimate of density.

It is possible to truncate the sightings at the edges of the unit shape. However this will make estimation of the detection function complex and will not use data from dropped sightings within the sampling circle but outside the edge of the unit. Thus, in this case, it is preferable to allow the circles to overlap rather than truncate at the unit edge. Typically the best approach is to either select a shape size so that the circle fits within it or to truncate the point transect data such that the circle fits within the shape.

Many point transect surveys are multi-species surveys, where the truncation radius is likely to vary by species. Thus, dependent on the unit size, the plots for some species may extend beyond the edges of a unit, whilst others will fall inside. In this case it may be preferable to select a unit size so that all plot sizes fit within a unit.

2.3.1 Triangular Units

A triangular unit, with the neighbourhood defined as the three adjoining units, combines a low number of edge units with good potential spread to detect clusters. However on the downside there tend to be large gaps between the circles and so potential observations are not maximised. This could be avoided by truncating the observations at the edges of the triangle rather than the circle, which as already described, will duplicate survey effort on overlapping areas and complicate the survey procedure by requiring extremely accurate radial angles to be measured.

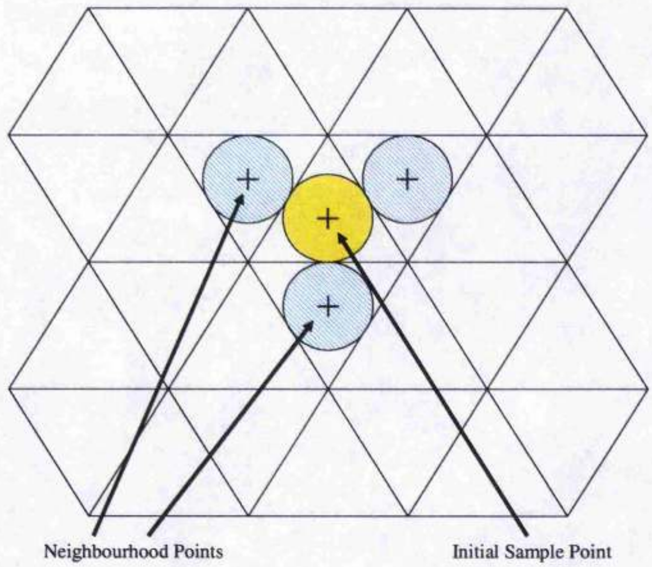


Figure 2.9: Triangle units combine a low number of edge units with good spread to detect clusters at the expense of comparatively large gaps between plots.

2.3.2 Square Units

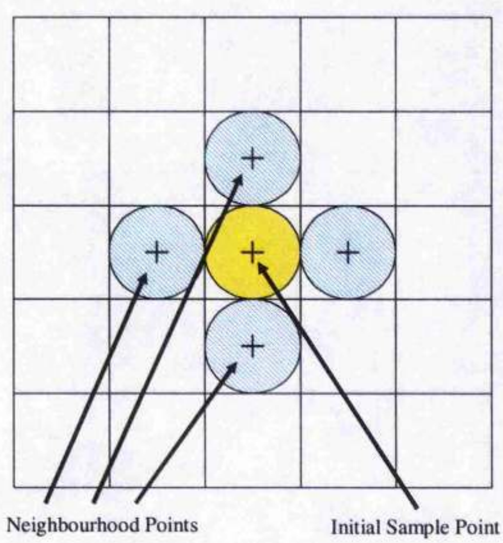


Figure 2.10: Square units. Thompson typically uses the north, south, east and west units as a neighbourhood, although many other combinations are possible.

A grid of squares provides a better fit to the circle than the triangle, and thus provides better coverage if the plot is to be kept within the boundary. As previously mentioned, Thompson typically uses a square grid with a cross shaped neighbourhood, in his examples. The cross pattern offers reasonable directional control, by allowing the network to spread following a cluster north, south, east or west. The ability to follow a cluster could be improved by adding the diagonally located corner squares, but this would be at the expense of potentially decreased efficiency from a larger number of edge units, and would only be of real benefit with a very highly clustered population.

2.3.3 Hexagonal Units

Circles fit well inside a hexagon, providing good coverage. Alternatively if the hexagon is fitted within the plot then, although the surveying procedures are complicated, the amount of overlap can be kept low and duplicated effort reduced. The use of an encircling neighbourhood, where all adjoining units are included, provides good compromise allowing the network to trace a cluster in all directions, with a slightly increased number of edge units compared with the cross pattern on a square grid. An alternative would be to use every other adjoining unit, so there are three units in the neighbourhood. In this case, when the plot is kept within the unit shape, the overall pattern is the same as for the triangular sampling units.

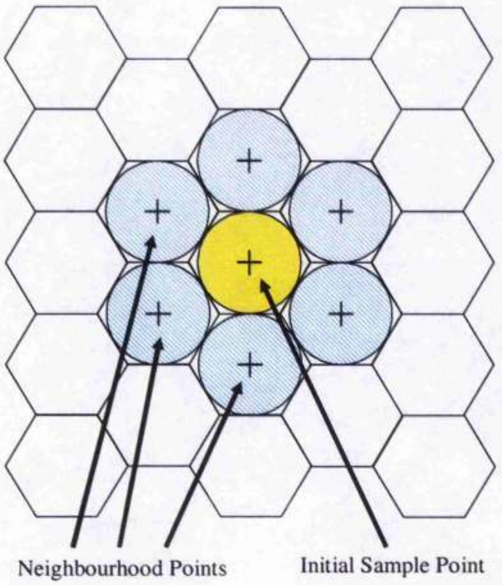


Figure 2.11: Hexagonal units provide a good spread to detect clusters, with minimal gaps between circles, although there are a comparatively high number of edge units.

2.3.4 Non-Contiguous Patterns

Thompson's methods do not require the neighbourhood units to be immediately adjacent to one another. As previously stated, the requirement is for symmetry of inclusion, such that the selection of any unit within a network will also include all other units within the same network. Figure 2.12 shows a grid of hexagons overlaid on the survey region. In this case the neighbourhood is defined as the three units in an equilateral triangle pointing down and spaced one hexagon away from the initially sampled unit. There are nine networks, eight of which are a single initial unit with no detections, and one consisting of an initial unit and three adaptive units. Network units are again shown with a thick black line as a boundary, and units that are part of the same network are shown with interconnecting thick black lines.

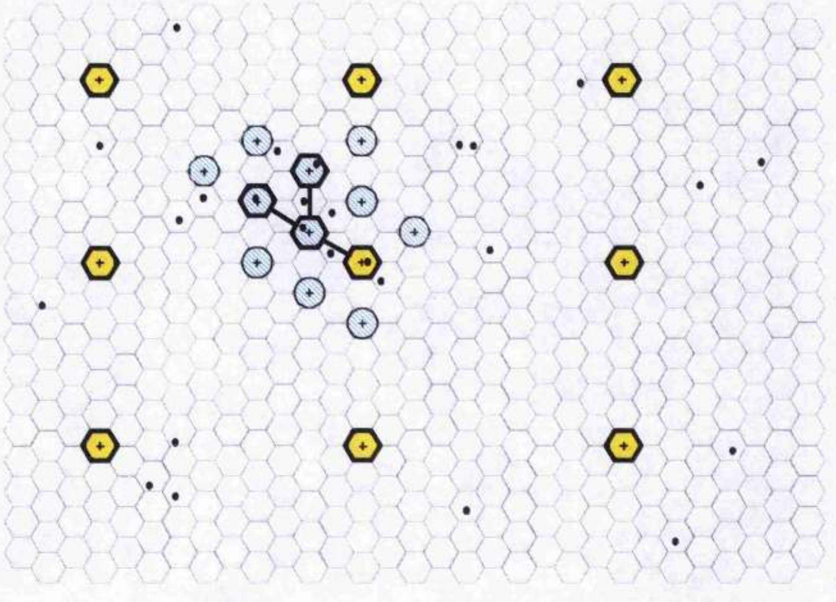


Figure 2.12: Example non-contiguous pattern on a grid of hexagons. Initial sample points are shown as solid (yellow) circles and the adaptive points as shaded circles. Networks have a thick black outline, where there is more than one plot in a network they are interconnected by a thick black line.

There are a number of occasions on which the use of non contiguous units will be of benefit. If the animals display responsive movement during the survey, either attraction to the point or avoidance, then it may be possible to reduce the impact by not adding the immediately adjacent neighbourhood points. This will depend on the behaviour of the survey species and the spacing, as there is a danger that this approach may cause even greater disturbance of the animals. It also depends on the clusters being sufficiently large, in relation to the truncation radius of the plot. It could also be useful as a mechanism to reduce the additional adaptive effort where clusters are large and so if you were to adapt to all adjacent points a significant amount of additional effort would be used. Finally if the clustering of the animals was such that you would expect them to be more widely spaced than twice the truncation radius it may then be preferable to provide gaps in the neighbourhood. A similar effect can be obtained by making the grid units larger than the surveyed plot so that the point is centred in the unit and truncation radius does not reach to the unit edges.

2.3.5 Grid Design Selection and Application

Each survey will need to be evaluated with reference to its particular characteristics. Ideally some level of simulation can be performed to identify the optimum grid and neighbourhood pattern, preferably using a pilot survey, previous survey data, or known species characteristics as input.

In the absence of specific information, and in considering the three basic patterns discussed so far, it is likely that a triangular or hexagonal distribution of units will, as a rule of thumb, provide the best option. Where there is extremely high clustering then the hexagonal pattern will be preferable as it allows the pattern to detect and follow a cluster in the most directions. Where the clustering is low, but sufficient to warrant an adaptive approach, then the triangular pattern is likely to be optimum.

Most survey regions do not have a regularly shaped contour and so there is likely to be some kind of edge effect where the units touch the edge. Strindberg (2001) investigates survey coverage and these techniques should be extensible to these methods. A conventional point transect survey is also likely to suffer from similar edge effects and so standard approaches can also be employed. If lines of points are used so that, in the definition of this chapter, the initial sample is systematically positioned, this is likely to be more difficult to fit to an irregular shape; and the calculations, such as the maximum number of primary units, more complex. The suggested approach would be to keep the lines the same length, all running at fixed parallel distances in the same direction and then randomise the start location of each line. For simplicity it will be easier to treat the points as randomly located, thus a more straightforward approach is to create a grid of points, which is randomly located over the survey region; where the grid consists of lines of points with equidistant (or approximately equidistant) spacing both between points in a line and the lines themselves. Analysis can then be performed using the Horvitz-Thompson-based estimators for a RIS design (equations 2.13 to 2.16).

It is suggested that units that overlap the edge of the survey region contour are excluded from the survey as, unless the area is extremely patchy or the plots are large compared to the survey region, this would only introduce small bias. Where initial points are on the edge of an area, so that neighbouring points would cross a boundary, then you do not include these points, although other points within the same neighbourhood that do fall inside the survey region will be included.

Alternatively, if the edge effect is thought to present an issue, a boundary can be applied to the grid of points so that the boundary fits within the survey region contour. A buffer zone, with a width of the truncation radius R, is then defined around the edge of the boundary. Only observations which lie within the boundary are included in the analysis, irrespective of whether the centre of the associated point transect lies within the boundary or the buffer zone. However effort for all point transects where the centre lies in the buffer zone is discounted. For analysis the observations from these points are added to the rest of the network within the boundary if one exists, otherwise, they can be added to nearby points that are themselves a network of size 1.

For most surveys the aim will be to maximise coverage and so the distance between points will have to be at least twice the truncation radius, to prevent overlapping sampling circles. Where the radius is not well known or understood, it will be sensible to increase the distance say to 2.5 or 3 times the estimated value. In these situations, simulation prior to surveying, using expected survey characteristics, will be extremely useful in helping to identify the optimum configuration to use.

2.4 Simulation

To illustrate the methods, a basic simulation was performed using the program RATS, described in Appendix A. A population was created in a square area of 130 units by 130 units, using a Poisson cluster process (Diggle, 1983). The number of parent clusters followed a Poisson(15) distribution, and the number of objects in each parent a Poisson (40) distribution. The vertical and horizontal coordinates of the centre of each parent cluster was selected from a Uniform (0, 130) distribution. Finally the position of each object within a parent cluster, relative to the cluster centre, was generated using a radial distance following a Normal(0, 4) distribution and an angle following a Uniform(0, 2π) distribution. For this example the group size was set to one; that is each object represented a single animal only. In the simulation used later in the thesis, this type of population is classified as a highly clustered population.

A single population was created containing 722 objects. A conventional and an adaptive point transect survey were then simulated, using 100 initial points. The points were equal spaced in a 10 by 10 grid at 12 unit intervals. A 5 unit boundary was applied all the way round the inside of the population area, to reduce edge effects, so that coordinates of the initial points were restricted to the range 5 to 125 both horizontally and vertically. The grid was then randomly located to fit within this inner boundary using a Uniform(0, 10) to get the horizontal and vertical offset.

The detection function was simulated using a half-normal key function with no adjustment terms, parameter $\sigma = 1.0$ and the truncation radius set to 2.0.

For the adaptive survey, the unit pattern was a simple grid of squares and the neighbourhood was defined using the standard Thompson convention of the four adjacent units above, below, left and right of the triggering unit. The squares were each 4 units by 4 units, so that the sampling circle of each point fitted exactly inside the square, and neighbouring units touched one another. Any adaptive points that overlapped the population edge were included, but no further adaptive points were added outside the edge. Thus there was a small edge effect, both from not continuing to follow a cluster with adaptive units, and that the points overlapping an edge only sampled a fraction of the sampling circle.

The adaptive survey added 82 adaptive points, making a total of 182 points surveyed. To provide a comparison a further conventional survey was run, but this time using a 14 by 14 grid of points on the same population, giving a total of 196 survey points. The overall grid dimensions remained at 120 units by 120 units, as for the 10 by 10 point grid, thus the inter-point spacing was reduced, and a new random location was used to position the grid. All other simulation parameters were kept the same.

Figure 2.13 shows the adaptive survey simulation, in which there were a total of 79 observations. The objects are shown as black dots, and any that have been detected are enclosed in a red square. The centre of each point is indicated by a cross and the circles represent the sampling area up to the truncation distance. All points that form a network are shown as solid (yellow) circles, whilst any edge units have diagonally hatched (blue) circles.

Although there are some large networks, careful inspection of the diagram reveals that no two initial points joined into a single network, although from the spread of networks this would not appear to be the case. So for example, if the large network towards the middle of the population is examined, it can be seen this emanates from the fifth initial point of the sixth row down. Although this network touches both the fourth initial point in the sixth row and the fourth and fifth initial points in the seventh row, none of these initial points have any detections and so each is actually a distinct network of size 1; whilst the network from the fifth initial point in the sixth row contains a total of 15 points.

Figure 2.14 shows the conventional survey using the same 100 initial points as the adaptive survey, with total of 7 observations. This is insufficient observations for a useful density estimate and in practice more points would be need to be surveyed. Figure 2.15 shows the conventional survey using 192 points, with the 14 by 14 grid of points located at a different random position. In this case there were 33 observations in total, which would also be typically insufficient for a reliable estimate. The simulation fixed the selection of estimated detection function to be the known function used in simulation, rather than pick from a choice of models. This will have removed some variance and allow for a more precise estimate than could normally be expected for this number of observations.

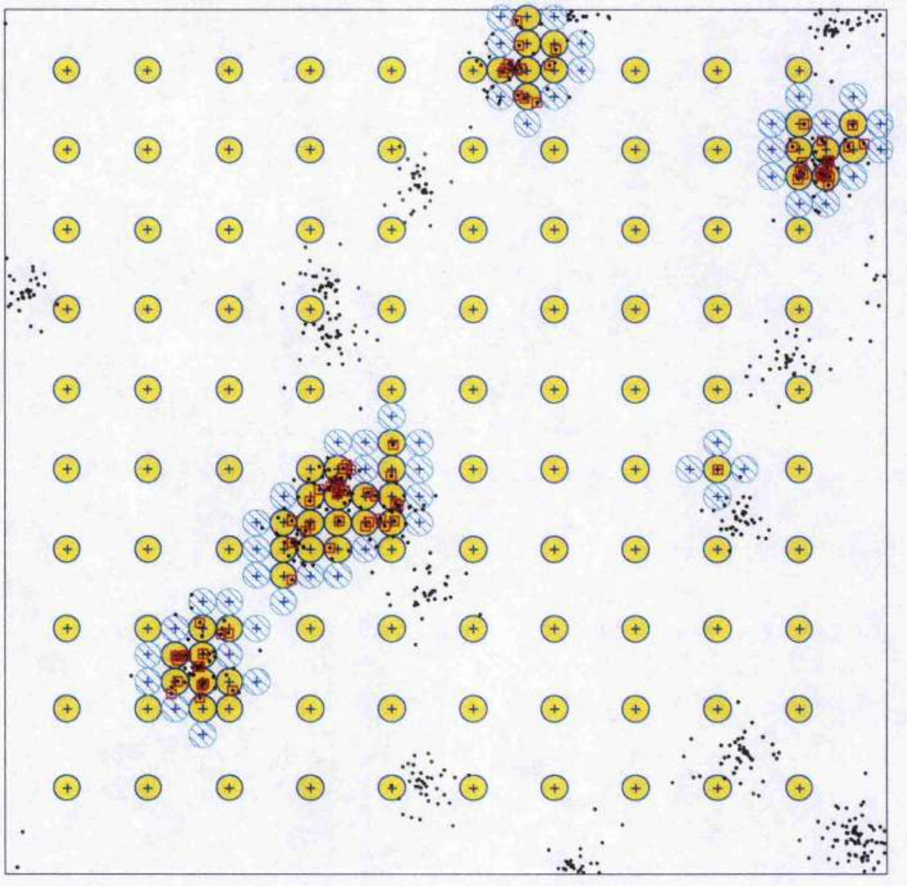


Figure 2.13: Simulation of an adaptive point transect survey on a highly clustered population. There were 100 initial sample points in a systematic grid and 82 adaptive points. The crosses identify the points and the circles the area covered at each point. The solid (yellow) circles identify points which belong to a network and the diagonally hatched (blue) circles are edge units. Objects are black dots and detected objects are surrounded by a red square.

The surveys used a randomly positioned, systematic grid of points, and is considered a RIS design. For each survey, the detection function was estimated by pooling all observations and using Distance 2.2 (Laake *et al.*, 1996) with model selection set to the half-normal model only. In each case no adjustment terms were added. The adaptive survey data were analysed with both the Hansen-Hurwitz and Horvitz-Thompson RIS based estimators, whilst the conventional surveys were analysed using the conventional distance formula.

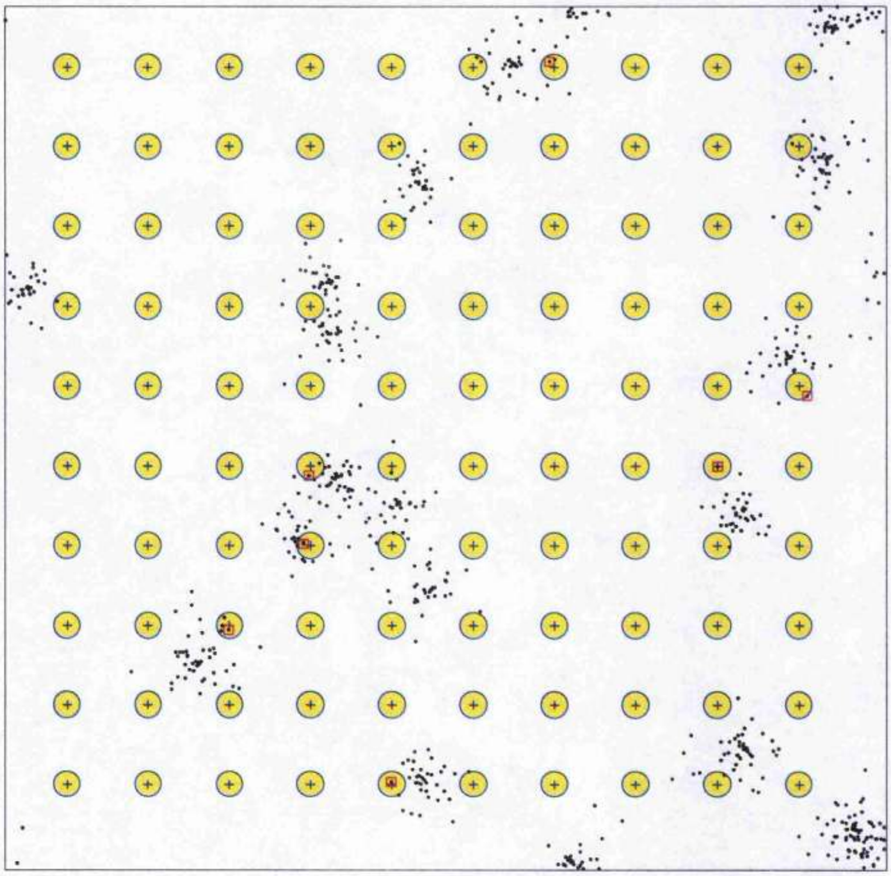


Figure 2.14: Conventional point transect survey, with 100 points in a systematic grid. The crosses identify the points and the solid (yellow) circles the area covered at each point. Objects are black dots and detected objects are surrounded by a red square.

The results are summarised in Table 2.4. To assist comparison, the numbers of objects in the conventional surveys have been converted to the mean numbers of observations per point. Upper and lower values for each density estimate have also been included, using a 95% log-normal confidence interval. There were 722 objects in the overall 130 by 130 unit frame, with a few of the clusters located on the edges. There was a 5 unit boundary on all edges and the survey grid had been randomly located within this inner 120 by 120 unit square, which contained only 540 objects, a lower density than the overall frame. Thus the true density, for the surveyed area, has been approximated as 540/(120x120) = 0.038. This is not strictly correct as the adaptive units were allowed to step outside the 120 x 120 unit frame and sample in the boundary area, although sampling stopped past the outside edge of the overall frame. This edge effect will be small and has been ignored for this basic simulation. The actual value of $h(0)$ used for simulation was 1.157, giving an expected mean number of objects per point of 0.204.

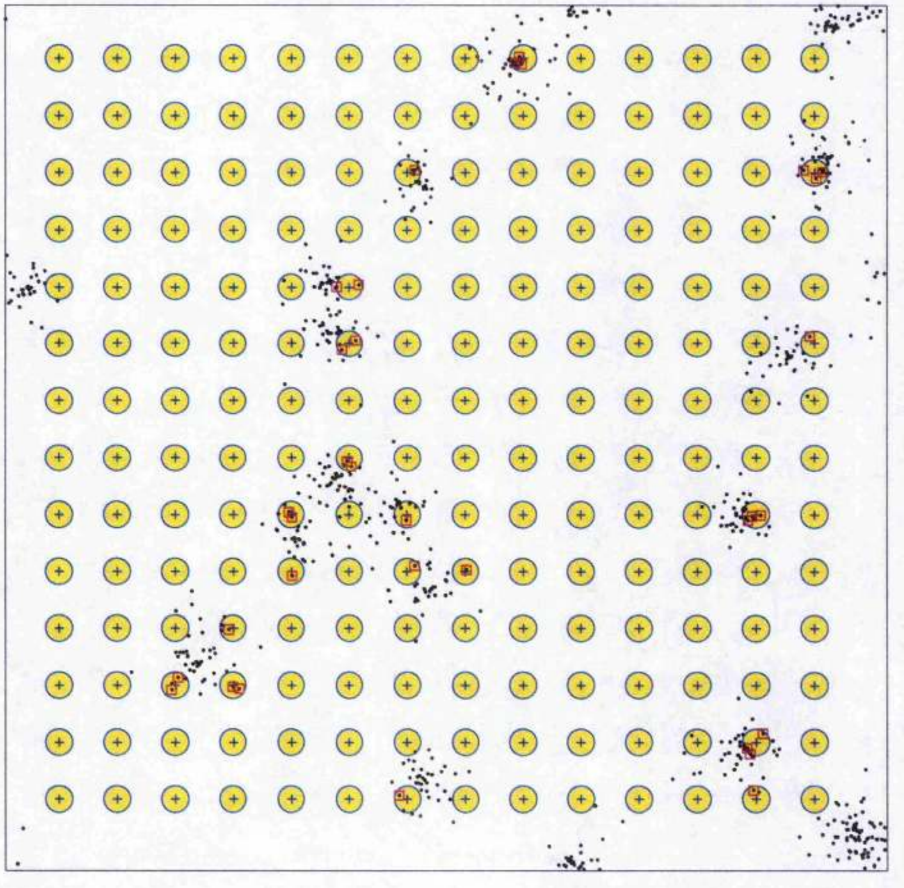


Figure 2.15: Conventional point transect survey, with 196 points in a systematic grid. The crosses identify the points and the solid (yellow) circles the area covered at each point. Objects are black dots and detected objects are surrounded by a red square.

All four density estimates underestimated the true density, but the true density lay inside the 95% confidence interval in all cases. There was a slight adverse edge effect in the network surrounding the seventh initial point in the top row of the adaptive survey, which would have contributed to the adaptive underestimate, but this would have been small.

The $h(0)$ estimate benefited from the adaptive approach with over two and half times the number of sightings of the Conventional 14 x 14 survey, reducing the *cv* from

23% for the Conventional 14 x 14 survey to 14% for the Adaptive survey. The *cv*s for the Conventional 14 x 14 survey and the Adaptive survey Horvitz-Thompson estimates of $E(\bar{n})$ were similar at 24% and 28% respectively, whilst the Hansen-Hurwitz was larger at 43%. Overall the Conventional 10x10 density estimate had a *cv* of 58%; the Conventional 14 x 14 a *cv* of 34%; the Adaptive Hansen-Hurwitz a *cv* of 46%; and the Adaptive Horvitz-Thompson 31%.

Table 2.4: Simulation results for conventional and adaptive simulations of the population of 722 objects.

	Conventional 10x10	Conventional 14 x 14	Adaptive Hansen-Hurwitz	Adaptive Horvitz-Thompson
Total Points	100	196	182	182
Total Detections	7	33	79	79
$\hat{h}(0)$	1.399	1.136	1.087	1.087
$\hat{V}(\hat{h}(0))$	0.3911	0.0698	0.0242	0.0242
$\hat{E}(\bar{n})$	0.070	0.168	0.096	0.143
$\hat{V}(\bar{n})$	0.000658	0.00166	0.00174	0.00159
$\hat{D}$	0.0156	0.0304	0.0167	0.0247
$\hat{V}(\hat{D})$	0.0000811	0.000104	0.0000579	0.0000599
$\hat{D}_{LCL}$	0.0054	0.0160	0.0071	0.0136
$\hat{D}_{UCL}$	0.0446	0.0577	0.0391	0.0450

2.5 Extensions

2.5.1 Secondary Species

Point transect surveys are typically multi-species so the trigger condition needs to be given careful consideration. If a rare species is important to the survey and it is expected to be highly clustered then this should be selected as the primary species and an appropriate trigger set based on the count of this species. The issue here is the prospect of reducing the efficiency of the secondary species estimates, as these will need to be estimated in the same way as the primary species. If it is known that there is no correlation between the primary species and a secondary species then the points could be potentially treated as totally random for the secondary species. However this is unlikely to be the case as there are many factors such as habitat, food supply, predators, etc., that may be correlated between the species.

A second policy would be to use a combination trigger. This could be as simple as the total count of observations at the point across all species; a total count across a subset of species, for example three particularly rare species; or a specific value for each species with say 2 sightings of species A, 4 sightings of species B or 1 sighting of species C all being valid trigger conditions.

2.5.2 Bootstrapping

Bootstrapping is frequently used to estimate variance and confidence intervals with distance sampling. An outline approach for point transects is provided in Buckland *et al.* (2001: p161). The same approach can be used with the adaptive methods of this chapter to estimate density precision. So for points, considered randomly located, the bootstrap sample is created by sampling with replacement from the k points in the initial sample and then including any associated networks for the resultant sample. For points systematically distributed on widely spaced lines then the resample unit will need to be the lines of initial points with the inclusion of any associated networks. In this case, as with a conventional point transect sample, the lines of points are sampled with replacement until the total number of initial points is the same or as close as possible to the number of initial points in the actual survey.

2.6 Discussion

In line transect sampling the prime component of the density estimate variance is commonly due to variance in n, whilst in point transect sampling the precision of $\hat{h}(0)$ can be poor and be the major contributor. Thus the basic adaptive approach outlined in this chapter provides a mechanism to increase the number of sightings and hence improve the precision of the detection function estimate. This is borne out by the simple simulation, where for a highly clustered population, the adaptive approach more than doubled the number of observations over those detected in a conventional survey with a similar total number of points surveyed. It is too early without extensive simulation to predict overall efficiency gains and at what level of clustering an adaptive approach benefits. In Chapter 4, more detailed simulation is carried out across three types of population providing more insight into the effects of clustering. However it is expected that adaptive point transect sampling will be mainly of use were the species is rare and a typical survey yields insufficient sightings for a reliable $h(0)$ estimate.

This adaptive approach introduces a degree of complexity which must be weighed against any advantage. The efficiency of the method will be highly dependent on the degree of clustering within the underlying population, and at this time suitable measures of just how clustered a population needs to be, are not readily available. In addition the distribution of initial points, neighbourhood pattern and trigger condition all need to be identified on a survey by survey basis. This is complicated by the fact that many point transect samples are multi-species, so the trigger may not just be the number of observations exceeding some limit for a single species. The truncation distance will be a major factor in the spacing of neighbourhoods and in the case of multi-species, where the optimal truncation radius may differ between species, it will be advisable to use twice the largest truncation radius as the minimal distance between points. Alternatively the truncation radius could be reduced to a less than optimal size for the species for which there were abundant sightings.

Until these design areas are developed further it is suggested that simulation should be used prior to any survey to identify the appropriate parameters to use. Ideally a pilot survey, previous survey data or known characteristics of the survey species should be used as input to the simulation. In particular the simulation should aim to estimate the number of adaptive points that will be added so that the survey can be adequately planned with the correct amount of resources available. If the adaptive behaviour leads to such a significant increase of points that the survey cannot be completed within the time available, be that due to monetary, biological or practical considerations, then there will be incomplete coverage and any benefit is likely to be lost.

Without proper planning the potential for getting things wrong is significant as the adaptive method does not limit the effort expended. At the most basic if the trigger is set too high the survey will not adapt and so you may not get the desired precision. In this case the trigger could then be revised down mid survey. Depending on survey procedure it may then be acceptable to revisit the neighbourhoods of points already surveyed that previously did not meet the trigger condition, but now do exceed the revised value. If the trigger is set too low then the adapting may be so excessive that the survey is no longer viable, as it will take too much effort to complete. The trigger can be revised up mid survey, but in this case you would need to drop any previously included neighbourhood points where the number of observations no longer meets the revised value. These and other approaches to limiting the total effort required for an adaptive survey are discussed in more detail in Chapter 3, Adaptive Line Transect Sampling.

Sampling neighbourhood points which by their very nature are in close proximity may induce bias through responsive behaviour, with either the attraction or evasion of animals. The additional adaptive points will also add to the off-effort travel for the survey; however the adaptive neighbourhood points will typically be nearer to the triggering point than the distance between the initial survey points. Thus the ratio of off-effort travel per point surveyed would actually be expected to decrease with an adaptive survey.

In the absence of any guidelines, then a triangular pattern is proposed using a neighbourhood of the three adjacent triangles. This provides the lowest number of edge units with the optimal cover for any related cluster of animals, and so minimises additional effort. If there is known to be a very high degree of clustering and the species is rare, then the hexagonal pattern can be considered with all adjacent hexagons forming the neighbourhood. The square or hexagonal pattern could also be considered if the addition of points to the survey has a low overhead in terms of cost or time and so is not a major concern.

The trigger will need to be based on at least some biological understanding of the species, however failing that a value of one provides a starting point. If the survey starts and it is quickly identified that the trigger is too low, then it can be raised as already described. For multi-species the trigger can be based on the rarest species.

It is not envisioned that applying the technique will introduce any significant field issues, above those already encountered with conventional point transect sampling. The major differences are the need to accurately locate the centre of points, particularly neighbouring points; to be wary of double counting points in neighbouring sampling circles; and to ensure there is minimal disturbance of animals when moving to neighbouring points.

With modern GPS equipment identifying the location of points should be practical unless the spacing is small, in which case a reel tape measure and a compass may be sufficient, although you may need to be careful not to disturb the area too much. A simple spreadsheet or computer program will probably be required to calculate the position of the neighbouring points. Depending on the equipment available and the type of area being surveyed this may need to be done prior to starting the survey. Thus all possible neighbourhood points could be pre-calculated and recorded on a map or in a table, perhaps to limit complexity only to an agreed depth of adaptation around each initial point (e.g. all combinations of 3 levels deep of triggering from an initial point).

When points are in close proximity there is a danger that animals near the edge of the truncation radius may be counted at both circles although they only lie within the truncation distance of one of the circles. Observers will need to be made aware of the dangers of this type of double counting, and should be given appropriate training, as with any distance sampling survey, prior to commencing. Practical training is invaluable, so items could be located at known distances and observers asked to estimate the distance and then provided with feedback to help them improve their estimates. It may also be feasible to place markers at the boundary of each point and its neighbour to minimise double counting of animals, although this is likely to lead to a degree of disturbance and so may be impractical.

To minimise disturbance, observers will need to take care when moving between points and typically wait some pre-agreed time before commencing the sample at each point to allow the environment to stabilise. It may also be possible to pre-mark out the initial points and some neighbourhood prior to the survey to prevent intrusion when locating the positions, particularly if an area is being re-surveyed using the same points year on year.

The approach here would not be applicable to cue counting, which can be addressed by employing the methods described for line transect sampling in Chapter 3. In some cases an adaptive approach may be suitable for trapping webs, although this is limited. In a trapping web, traps are aligned in circles on radiating lines from the centre of a web. Thus for trapping webs a detection is actually the trapping of an animal, and you estimate the probability density function of trapping distances from the centre of the web, so that each web represents a point. In the past trapping was usually repeated using the same web over a series of days or nights, and thus there is no obvious analogy for the neighbourhood as it has been employed in this chapter. Current good survey practice proposes the use of a large number of separate webs, in which case with sufficient locations the method could be adopted. As the cost of the web can be relatively high, webs are often re-used to cover the whole area, and the separate webs are not all surveyed simultaneously. Given a suitable number of webs then an adaptive approach could be considered, so long as all the locations can be trapped within the available time. In addition the cost to implement each web is likely to be relatively high, so if the number of webs is low compared to the number of units added for each neighbourhood, it would probably only be justified for rare populations with a high degree of clustering.

In the Chapter 3 we proceed to develop Thompson's adaptive estimators for line transect sampling and address some of the issues specific to this as well as exploring ways in which the total effort can be limited for either lines or points.

Chapter 3

Adaptive Line Transect Sampling

3.1 Introduction

In this chapter we continue the application of Thompson's methods to distance sampling, but this time focussing on line transect sampling.

Line transect sampling (Buckland *et al.*, 2001) is an extension of strip transect sampling (see for example, Thompson, 1992; Buckland *et al.*, 2001; Seber, 1982); however it does not require all objects to be detected in the strip, and thus typically allows a wider strip to be used and for more observations to be recorded. The method is widely used for estimating wildlife abundance, and in particular for marine surveys of cetaceans. Surveying can be from land, sea or the air and in some cases a combination of these is used. For example the SCANS survey (Hammond *et al.*, 1995) used both ships and aircraft to survey harbour porpoise and other small cetaceans.

In line transect sampling, the observer follows a series of straight lines, *tracklines*, recording observations and their perpendicular distance from the trackline. An observation, or detection, of an object may relate to a single animal or a group of animals, in which case the distance is estimated to the centre of the group. The perpendicular distance may be estimated directly, or recorded as a radial distance and angle, relative to the trackline, from which the perpendicular distance is calculated. These distances are used to estimate a detection function, $g(x)$, which is the probability of detecting an object at distance x from the trackline. This is then used to estimate an *effective strip half-width*, which is a measure of distance from the trackline to the edge of a notional strip. The area of this notional strip represents a strip which, assuming all objects were detected, would produce the same count of detections as was actually recorded. This is analogous to the effective area estimated for point transects.

Line transect surveys will typically consist of long lines located either randomly or more commonly systematically over the survey area, and the estimators developed here allow for both types of design. For the examples in this chapter we will typically add short adaptive segments on either side of the main transect, which by the definition of section 3.1.1, Survey Designs, is a systematic initial sample (SIS) design. This will be more appropriate for many applications of line transect sampling, and so we go on to extend the estimator using a Horvitz-Thompson-based estimator for a SIS design.

This chapter follows the same format as the adaptive point transect chapter, and builds on the approach as before. Thus it starts by defining the basic line transect estimating equations and merges these with adaptive sampling. It proceeds to define four basic estimators using Thompson's core Hansen-Hurwitz and Horvitz-Thompson-based adaptive estimators, applied to both random initial sample (RIS) and SIS survey designs. Although much of the notation is similar, there are some small differences, such as the switch from points to transects and transect legs. Thus the basic estimators are repeated here, with notation redefined as appropriate. However due to the similarity in which they operate we do not include the detailed worked examples of the previous chapter.

The probability of detection is typically a combination of many factors and not just distance from the trackline. We start this chapter by considering a purely distance-based detection function. Later in the chapter, as with the point transect chapter, we develop an adaptive density estimator that can include covariates in the detection function estimation. Example covariates would include factors such as vegetation cover, group size, time of day, sea state, wind, sun angle or any other variable that may affect sighting objects.

The chapter discusses appropriate adaptive patterns and survey designs and then proceeds to consider extensions to the methods. In particular it looks at approaches to limit the total adaptive effort; the principles of which can be applied to both adaptive point and line transect sampling.

As before we define the *survey region* as the area which contains the *population* of objects for which we are producing estimates. We use a grid of units, with transects or transect legs run through each unit. Depending on the truncation half-width the transect strip may have a greater or smaller area than the unit. If the transect strip area is greater than the unit then transects will potentially overlap. Each unit represents a potential transect and so the *surveyed area* is the total number of units in the grid multiplied by the area of each transect. Thus the surveyed area can be greater than the area of the survey region, this is not an issue and the estimators are design to cope with this. If the survey uses a vehicle such as a ship or an aircraft this is referred to as the *observation platform*.

As transects will not always be contiguous, effort is required for the observers or observation platform to move between transects or transect legs. This unproductive effort is referred to as *off-effort*.

Throughout the chapter, the terms object and animal group are used interchangeably as in the previous chapter.

3.1.1 Survey Designs

We consider transects as following the centreline of units within a grid which has been randomly located over the survey region. These units will normally be rectangular in shape and are expected to be long and thin. Adaptive sections will typically be added on either side of a triggering transect so that they are running parallel to the initial trackline.

For this chapter we again consider two types of survey design, a Random Initial Sample (RIS) and a Systematic Initial Sample (SIS). The definition is the same as for the point transect chapter, but is redefined here in terms of line transects.

For the RIS designs we consider that the trackline running through a unit in the grid represents a single transect. For a SIS design there are both primary and secondary units, so the trackline running through a unit in the grid is considered a transect leg; and the combination of a number of secondary units into a single primary unit forms a transect.

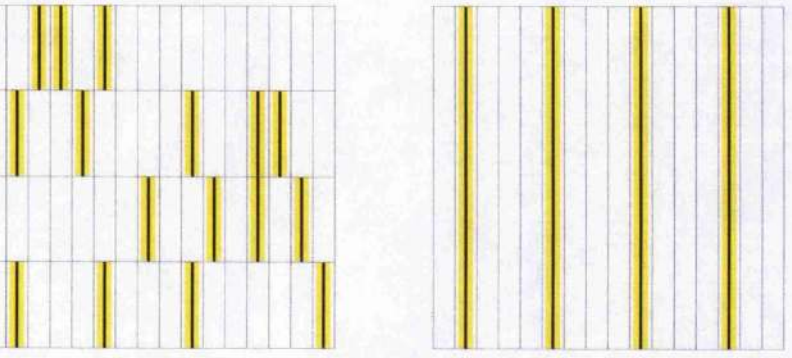


Figure 3.1: RIS line transect survey designs, with the transect lines shown as a thick black line and the surrounding sampling strip as solid (yellow) rectangles. In the left hand design transect lines have been selected at random for both vertical and horizontal location. In the right hand design the transects in the grid run the full height of the survey region, and are systematically spaced in the horizontal axis. Thus, given that the grid has been randomly located over the survey region, this is assumed to be a RIS design.

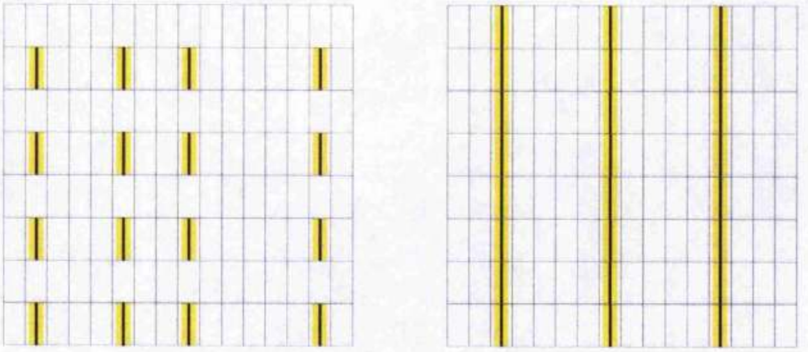


Figure 3.2: SIS line transect survey designs, with the transect lines shown as thick black lines and the sampling strip as solid (yellow) rectangles. In the left hand design transect legs are systematically distributed in lines (the primary units) which are then randomly (in the horizontal plane) located on the survey region. In the right hand design, transect legs are systematically located in lines (the primary units). However in this case there is no gap between transect legs and thus the 8 legs making up each overall transect appear as a single transect. The transects (primary units) are then systematically distributed over the survey region, making this a SIS design.

If transects are selected at random from the grid, then this is clearly considered a RIS design. However if the grid itself is randomly located on the survey region, with transects either systematically or randomly spaced within it, then it is accepted distance sampling practice to also consider this a RIS design (Figure 3.1), with the result that variance will be slightly overestimated.

If the primary unit spacing is significantly different from the secondary unit spacing, then the design is considered a systematic sample as shown by the SIS designs in Figure 3.2.

3.2 Theory

As with point transects we have four basic estimators to consider: RIS and SIS survey designs, combined with Hansen-Hurwitz and Horvitz-Thompson-based estimates. However when considering extensions to the methods, we focus on the Horvitz-Thompson-based estimator for a SIS design.

3.2.1 Line Transect Sampling Basic Formulae

From Buckland *et al* (2001: p54) the density of a population from line transect sampling is given by

$$D = \frac{E(n) \cdot f(0) \cdot E(s)}{2L}$$

where

$E(n)$	is the expected number of animals in the surveyed area
$f(0)$	is the value of the probability density function $f(x)$ of distances to detected objects, evaluated at $x = 0$
$E(s)$	is the expected group size for the population
L	is the total effort used foe the survey (the total transect length)

Replacing parameters by their estimators then an estimate of the density is given by

$$\hat{D} = \frac{n \cdot \hat{f}(0) \cdot \hat{E}(s)}{2L} \tag{3.1}$$

Using the delta method (Seber, 1982: p5-7) the variance can be estimated by

$$\hat{\text{var}}\left(\hat{D}\right) = \hat{D}^2 \cdot \left[\frac{\hat{\text{var}}[n]}{n^2} + \frac{\hat{\text{var}}\left[\hat{f}(0)\right]}{\left[\hat{f}(0)\right]^2} + \frac{\hat{\text{var}}\left[\hat{E}(s)\right]}{\left[\hat{E}(s)\right]^2}\right] \tag{3.2}$$

3.2.2 Merging Approaches

The basic approach is to overlay the survey region with a grid of units, with transect tracklines conducted through any selected units. Adaptive units are then added in the neighbourhood of any units which meet the adaptive trigger condition, which will typically be based on the number of observations in the unit.

As with adaptive point transect sampling in Chapter 2, Thompson's methods are used to estimate the number of observations and group size, whilst the overall $f(0)$ estimate is made by pooling all observations. As before this assumes no heterogeneity is introduced by the adaptive process. We also continue the assumption that the probability of detection on the line is certain, $g_0 = 1$; use a previously known value; or estimate it using an alternative mechanism.

In section 2.2.2 it was shown that Thompson's underlying adaptive estimators can be applied to point transect sampling, even though Thompson's methods are primarily founded on the principle that all objects in a sampling unit are detected, whilst distance sampling is based on the premise that only a proportion of the objects are detected. The same argument can be directly applied to line transect sampling, and thus Thompson's core adaptive estimators (equations 1.1 to 1.8) can be used in their basic form.

The units in the grid will typically be rectangular, as it is an ideal fit with the rectangular shape of the sampling strip; although alternative unit shapes are feasible, they are not explored further in this thesis. The transect trackline is run down the centreline of the rectangular grid unit (Figure 3.3).

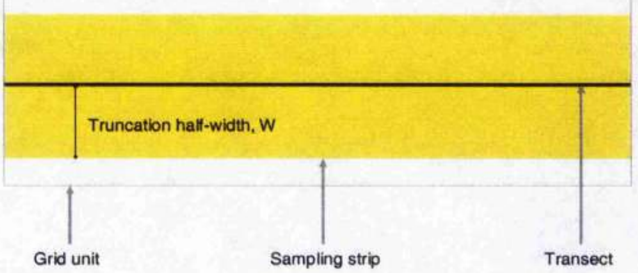


Figure 3.3: The transect trackline is run through the centre of the rectangular grid unit. The transect is shown as a thick black line and the sampling strip, with truncation half-width, *W*, is shown as a solid (yellow) rectangle.

The grid itself is randomly located on the survey area, with transects either systematically or randomly spaced. Each unit is treated as a separate sampling unit for the analysis using Thompson's methods.

To simplify survey procedures the truncation-width will typically be such that the sampling strip is contained within the unit, although there is no mathematical requirement for this. The field implications of varying strip widths are discussed further in the section 3.3.

For the RIS and SIS survey designs considered in this section we use the simple neighbourhood pattern of the two units, one each side, of the unit being sampled, so that the adaptive transect tracklines run parallel with the triggering transect. For the RIS designs, each unit is considered a separate transect. For the SIS designs, which use primary and secondary units, each grid unit (secondary unit) is considered a transect leg; and the primary units consist of a number of secondary units and so are considered transects. This is a notational issue only. In conventional line transect sampling, the number of transects is sometimes used in variance estimation, and so influences the resulting variance estimate. In this case it is the number of primary and secondary units that influence the variance estimate, and thus the naming does not affect the result. Each secondary unit can, if required, be referred to as a separate transect, however the use of transect legs is felt a preferable terminology.

Based on the definitions used with point transect sampling, then

N	is the number of individual animals in the survey region
δ	is the number of individual animals in the surveyed area
N_g	is the number of animal groups in the survey region
τ	is the number of animal groups in the surveyed area
K	is the number of units (transects for a RIS design and transect legs for a SIS design) in the survey region
A	is the total area for the survey region
A_s	is the surveyed area
l	is the effort used on a trackline through a single unit
L	is the total survey effort
W	is the truncation half-width

The surveyed area is the product of the sampling strip area in a unit, multiplied by the number of units in the grid, so

$$A_s = 2lKW \tag{3.3}$$

and repeating equations 2.3 and 2.4 we have

$$N = \frac{A}{A_s}\delta$$

and

$$N_g = \frac{A}{A_s}\tau$$

In the following two sections, we briefly illustrate the adaptive mechanism for a RIS and then a SIS design survey.

Example RIS Survey

In this example we use a grid consisting of rectangles running vertically, for the height of the survey region, with four initial transects. Although the initial units (transects) are systematically spaced in the horizontal axis; the grid has been randomly located over the survey region and so, in line with distance sampling conventions, we consider this a RIS design (Figure 3.4).

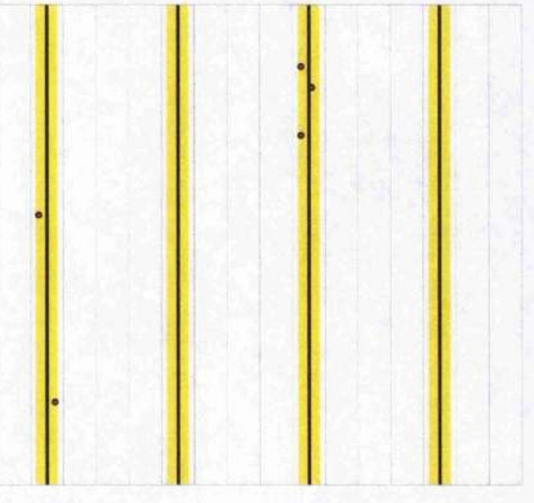


Figure 3.4: RIS survey example. A grid of rectangles is randomly located over the survey area. Transects are run down the centre line of each sampled unit (rectangle). There are four initial transects running vertically through the survey region. The transect tracklines are shown as a thick black line, the transect sampling strip is shown as a solid (yellow) rectangle and detections are shown as (red) dots.

The trigger condition is set at a single detection on a transect, and adaptive units are added either side of any initial transect which meets the condition. If any adaptive transects also meet the condition then adaptive transects are added to that. However as the neighbourhood is only the transect to either side, then an adaptive transect will only add one more transect, because on one side it will already have the triggering transect. The process continues until the trigger is no longer met; or the adaptive transects are limited by another transect or the edge of the survey region. In the example, two of the initial transects met the condition and a total of six adaptive transects were added (Figure 3.5).

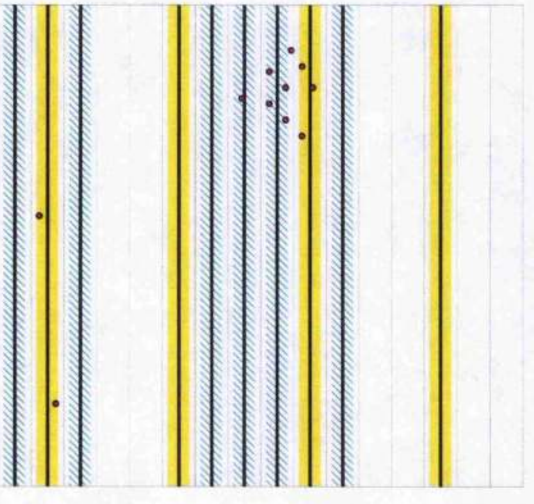


Figure 3.5: Adaptive transects are run in the parallel units to the left and right of any transect that meets the trigger condition. In this example the trigger condition is a single detection on a transect (sampling unit). The sampling strips for the adaptive transect are shown with (blue) cross-hatching, whilst the sampling strip for the initial legs are still shown as solid (yellow) rectangles.

Networks are then formed from neighbouring transects that meet the trigger condition, with the property that any transect that meets the conditions, would also include all other transects in the same network. This is referred to in this thesis as the *symmetry of inclusion*. Any initial unit that does not meet the condition is also considered a network of size 1. In the example two of the adaptive units form a network with the third initial transect from the left, and there are a total of four networks in the sample (Figure 3.6).

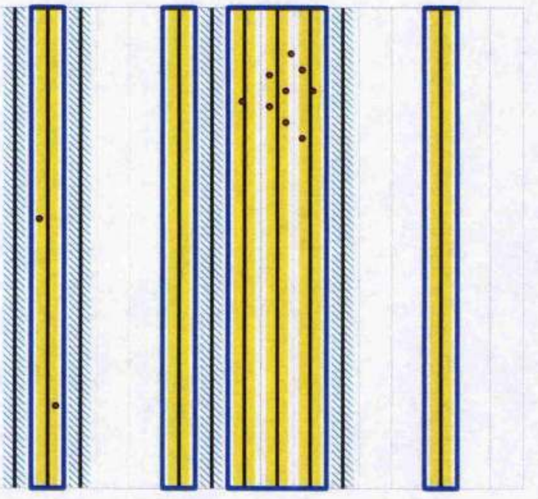


Figure 3.6: All adaptive legs have been added. There are four networks, which are enclosed in thick (blue) lines. Sampling strips for the transect are now shown as a solid (yellow) rectangle, if the transect belongs to a network; or with (blue) cross-hatching if the transect is an edge unit.

Example SIS Survey

In this example we run the transect tracklines horizontally, whilst in the previous RIS example they were run vertically. There is no significance in this, and the grid can be aligned in any direction, (although, following good distance sampling practice, if there is a density gradient in the population then transects should be aligned so they are parallel to any gradient).

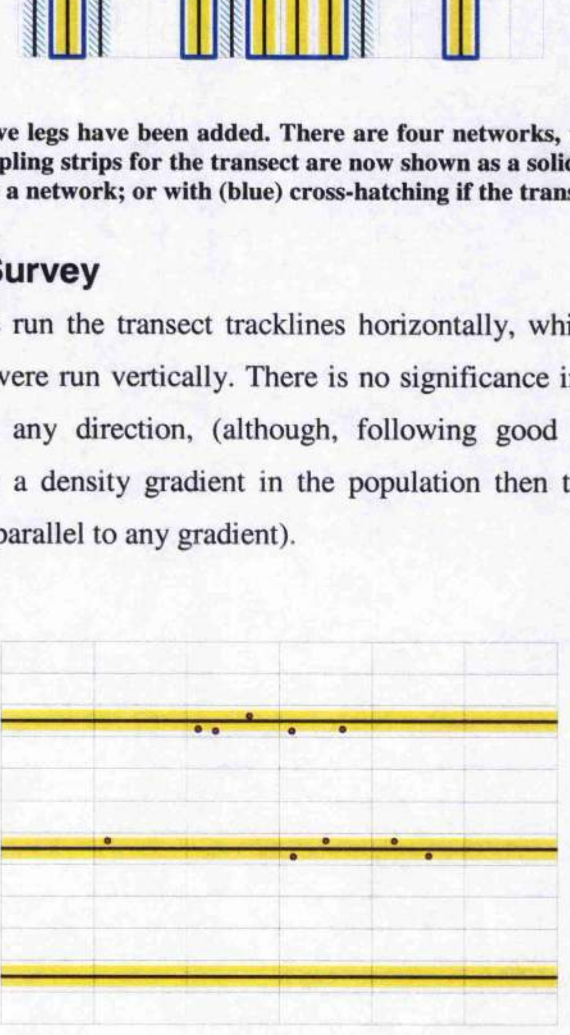


Figure 3.7: SIS survey example. A grid of rectangles is randomly located over the survey area. Transect legs are run down the centre line of each sampled unit (rectangle). There are three initial transect (primary units) running horizontally through the survey region. Each initial transect consists of 6 transect legs (secondary units), in a contiguous line. The transect legs tracklines are shown as a thick black line, the transect leg sampling strip is shown as a solid (yellow) rectangle and detections are shown as (red) dots.

The transects are the primary units and each consists of six transect legs (secondary units) which run contiguously, so that they join into a single trackline. There are three primary units, spaced systematically, and the grid has been located randomly over the survey region. If the spacing between transects was similar to the spacing between the transect legs in a primary unit, then it would have been reasonable to consider this a RIS design (in which case we would have considered each unit to contain a transect and not a transect leg). However the spacing is not similar and so we consider this a SIS design (Figure 3.7).

The neighbourhood is again the units on either side of a triggering unit, thus in this case adaptive transect legs are added above and below a transect leg that meet the condition. With the trigger set to one detection on a transect leg, twelve adaptive legs are added (Figure 3.8).

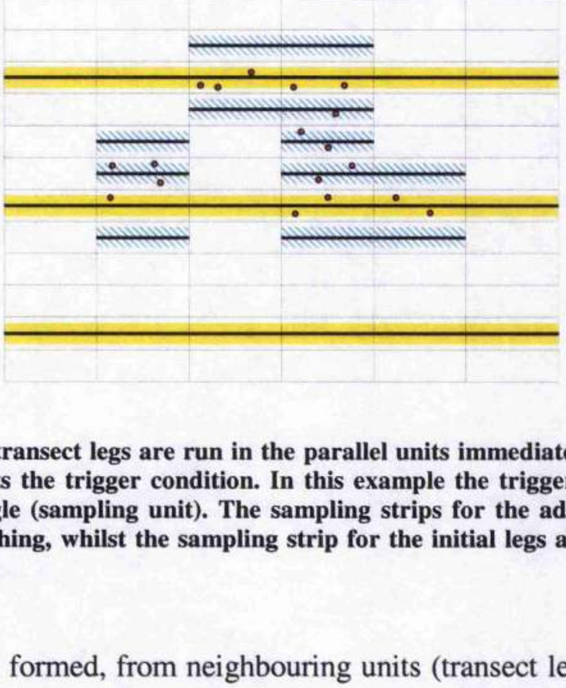


Figure 3.8: Adaptive transect legs are run in the parallel units immediately above and below a transect leg that meets the trigger condition. In this example the trigger condition is a single detection in a rectangle (sampling unit). The sampling strips for the adaptive legs are shown with (blue) cross-hatching, whilst the sampling strip for the initial legs are still shown as solid (yellow) rectangles.

Networks are again formed, from neighbouring units (transect legs) which meet the condition. There is a total of seventeen networks; thirteen of these are initial transect legs with no detections; four contain detections; and two of these form networks of more than a single unit. In one case the network has spread so as to link secondary units from different primary units in the same network (Figure 3.9).

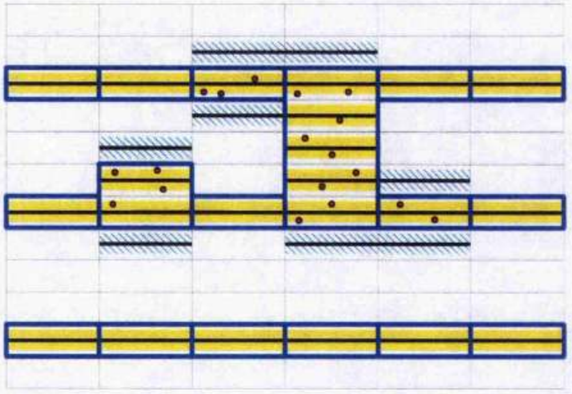


Figure 3.9: All adaptive legs have been added. There are seventeen networks, which are enclosed in thick (blue) lines. Only two are larger than one unit, and only four contain detections. Sampling strips for the transect legs are now shown as a solid (yellow) rectangle, if the leg belongs to a network; or with (blue) cross-hatching if the transect leg is an edge unit.

Expected Number of Observations

From section 2.2.2 we have

$$E(n) = E(\overline{n}) \cdot k_s$$

where

k_s is the number of transects in the total sample

so an estimate of the expected mean number of observations in the sample is

$$\hat{E}(n) = \hat{E}(\overline{n}) \cdot k_s$$

Expected Group Size

We start by estimating the expected group size by the mean observed group size, that is, as the total number of animals observed divided by the number of observations.

Let

a be the total number of animals observed in the total sample

a_j be the total number of animals observed on the j^{th} transect

a_{jX} be the number of animals in the X^{th} observation of the j^{th} transect

So

$$a_j = \sum_{X=1}^{n_j} a_{jX}$$

where

n_j is the number of detections on the j[th] transect

and

$$a = \sum_{j=1}^{k_s} a_j$$

where

k_s is the number of transects surveyed

Let $E(a)$ be the expected total number of animals observed had there been no adaptation in the survey and $E(\overline{a})$ the expected mean number of animals observed per transect. Then

$$E(a) = E(\overline{a}) \cdot k_s$$

so that an estimate of the expected number of animals observed in the sample is

$$\hat{E}(a) = \hat{E}(\overline{a}) \cdot k_s$$

and thus an estimate of the expected group size is

$$\hat{E}(s) = \frac{\hat{E}(a)}{\hat{E}(n)} = \frac{\hat{E}(\overline{a}) \cdot k_s}{\hat{E}(\overline{n}) \cdot k_s} = \frac{\hat{E}(\overline{a})}{\hat{E}(\overline{n})}$$

Density Estimate

Substituting the adaptive estimators of $E(\overline{n})$ and $E(s)$ in the conventional line transect density equations (3.1 and 3.2) gives:

$$\hat{D} = \frac{\hat{E}(n) \cdot \hat{f}(0) \cdot \hat{E}(s)}{2L} = \frac{\hat{E}(\overline{n}) \cdot \hat{f}(0) \cdot \hat{E}(s)}{2l}$$

where

$\hat{E}(n)$ is an unbiased estimate of the expected number of observations in the sample

$\hat{E}(\overline{n})$ is an unbiased estimate of the expected mean number of observations per transect

$\hat{f}(0)$ is an estimate of the of the probability density function $f(x)$ evaluated at $x = 0$

$\hat{E}(s)$ is an unbiased estimate of the expected group size for the population

L is the total effort

l is the effort used to sample a unit. So for a RIS designs this is a single transect, whilst for a SIS design it is a single transect leg

and

$$\hat{\mathrm{var}}(\hat{D}) = \hat{D}^2 \cdot \left[\frac{\hat{\mathrm{var}}\left[\hat{E}(\bar{n})\right]}{\left[\hat{E}(\bar{n})\right]^2} + \frac{\hat{\mathrm{var}}\left[\hat{f}(0)\right]}{\left[\hat{f}(0)\right]^2} + \frac{\hat{\mathrm{var}}\left[\hat{E}(\bar{s})\right]}{\left[\hat{E}(\bar{s})\right]^2} \right]$$

With the group size estimated by the mean observed group size then the estimate simplifies to

$$\hat{D} = \frac{\hat{E}(\bar{n}) \cdot \hat{f}(0) \cdot \hat{E}(s)}{2l} = \frac{\hat{E}(\bar{n}) \cdot \hat{f}(0) \cdot \hat{E}(\bar{a})}{2l \cdot \hat{E}(\bar{n})} = \frac{\hat{E}(\bar{a}) \cdot \hat{f}(0)}{2l} \tag{3.4}$$

and

$$\hat{\mathrm{var}}(\hat{D}) = \hat{D}^2 \cdot \left[\frac{\hat{\mathrm{var}}\left[\hat{E}(\bar{a})\right]}{\left[\hat{E}(\bar{a})\right]^2} + \frac{\hat{\mathrm{var}}\left[\hat{f}(0)\right]}{\left[\hat{f}(0)\right]^2} \right] \tag{3.5}$$

The density of animal groups (observations) is obtained by replacing the group size estimator, $\hat{E}(s)$ by 1 giving

$$\hat{D}_g = \frac{\hat{E}(\bar{n}) \cdot \hat{f}(0)}{2l} \tag{3.6}$$

and

$$\hat{\mathrm{var}}(\hat{D}_g) = \hat{D}_g{}^2 \cdot \left[\frac{\hat{\mathrm{var}}\left[\hat{E}(\bar{n})\right]}{\left[\hat{E}(\bar{n})\right]^2} + \frac{\hat{\mathrm{var}}\left[\hat{f}(0)\right]}{\left[\hat{f}(0)\right]^2} \right] \tag{3.7}$$

3.2.3 Assumptions

As with point transect sampling, the standard assumptions apply. For clarity these are re-iterated here:

(i) Probability of detection on the line $g(0)$, is 1, or suitable methods are used to estimate the actual value.

(ii) There is no size bias (the probability of detection is independent of group size).

(iii) There is no responsive movement of animals in advance of detection.

In addition the following additional assumption is made:

(iv) Probability of detection is independent of whether or not effort is adaptive. i.e. probability of detection is only a function of distance from the line and the adaptive procedure does not induce heterogeneity in the $f(0)$ estimate.

As was done with adaptive point transect sampling, later in this chapter we discuss approaches to remove or reduce the need for assumptions (i), (ii) and (iv).

3.2.4 Adaptive Line Transect Sampling Estimators for Mean Number of Groups Detected and Mean Number of Animals Detected

Here we define the same four basic estimators for $E(\overline{n})$ and $E(\overline{a})$, as in Chapter 2, but this time for line transects. Thompson has two underlying estimators formed from the Hansen-Hurwitz and the Horvitz-Thompson estimators and versions of these are presented for both RIS and SIS survey designs, giving a total of four estimators. The line transect estimators follow the same form as for point transects and are primarily unchanged although there are subtle changes in the notation. We also add a with replacement variance estimate for the RIS estimators.

Care needs to be taken in appreciating the difference in notation between RIS and SIS designs. For a RIS design each unit of the grid represents a transect and for a SIS design each unit represents a transect leg. Thus for a RIS design, networks are made up of transects, whilst for a SIS design, networks are made up of transect legs. So for a RIS design n_{ijX} relates to X^{th} observation of the j^{th} transect in the i^{th} network; whilst for a SIS design, each unit represents a transect leg and so n_{ijX} relates to the X^{th} observation of the j^{th} transect leg in the i^{th} network.

Hansen-Hurwitz-based Estimators

$E(\overline{n}_{HH})$ Estimate for RIS Design

As in equations 2.9 and 2.10 the line transect estimates of expected mean number of observations per transect are

$$\hat{E}(\overline{n}_{HH}) = \frac{1}{k}\sum_{i=1}^{k} w_i \tag{3.8}$$

where

$$w_i = \frac{1}{m_i}\sum_{j\in\psi_i} n_{ij}$$

k	is the number of transects in the initial sample
ψ_i	is the network which includes the ith initial transect
m_i	is the number of transects in the network ψ_i
n_{ij}	is the number of observations at the jth transect within the network ψ_i

The estimate of variance, assuming the initial sample is selected without replacement, is

$$\hat{V}\left(\hat{E}(\overline{n}_{HH})\right) = \frac{(K-k)}{Kk(k-1)}\sum_{i=1}^{k'}\left(w_i - \hat{E}(\overline{n}_{HH})\right)^2 \qquad (3.9)$$

where

K	is the total number of (potential) transects in the survey region

From Thompson (1996: p100), then if the initial sample is selected with replacement, the variance is estimated by

$$\hat{V}\left(\hat{E}(\overline{n}_{HH})\right) = \frac{1}{k(k-1)}\sum_{i=1}^{k}\left(w_i - \hat{E}(\overline{n}_{HH})\right)^2 \qquad (3.10)$$

$E(\overline{a}_{HH})$ Estimate for RIS Design

Similarly an estimate of the expected mean number of animals observed per transect is

$$\hat{E}(\overline{a}_{HH}) = \frac{1}{k}\sum_{i=1}^{k} u_i \qquad (3.11)$$

where

$$u_i = \frac{1}{m_i}\sum_{j\in\psi_i}\sum_{X=1}^{n_{ij}} a_{ijX}$$

a_{ijX}	is the number of animals in the Xth observation of the jth transect within the network ψ_i

The estimate of variance, assuming the initial sample is selected without replacement, is given by

$$\hat{V}\left(\hat{E}(\bar{a}_{HH})\right)=\frac{(K-k)}{Kk(k-1)}\sum_{i=1}^{k}\left(u_i-\hat{E}(\bar{a}_{HH})\right)^2 \tag{3.12}$$

If the initial sample is selected with replacement, then the variance is estimated by

$$\hat{V}\left(\hat{E}(\bar{a}_{HH})\right)=\frac{1}{k(k-1)}\sum_{i=1}^{k}\left(u_i-\hat{E}(\bar{a}_{HH})\right)^2 \tag{3.13}$$

$E(\bar{n}_{SIS.HH})$ Estimate for SIS Design

From the point transect estimates, equations 2.17 and 2.18, an estimate of the expected mean number of observations per transect leg is

$$\hat{E}(\bar{n}_{SIS.HH})=\frac{1}{r}\sum_{h=1}^{r}w_h \tag{3.14}$$

where

$$w_h=\frac{1}{M}\sum_{i=1}^{\kappa_h}\frac{\sum_{j\in\psi_i}n_{ij}}{t_i}$$

r	is the number of initial transects (primary units) in the sample
M	is the number of transect legs (secondary units) in each initial transect
κ_h	is the number of networks that intersect the hth initial transect
ψ_i	is the set of points in the ith network
n_{ij}	is the number of observations on jth transect leg of the ith network
t_i	is the number of initial transects that intersect the ith network

The variance is estimated by

$$\hat{V}\left(\hat{E}(\bar{n}_{SIS.HH})\right)=\frac{s_W^2}{r}\left(1-\frac{r}{R}\right) \tag{3.15}$$

where

$$s_W^2=\frac{1}{r-1}\sum_{h=1}^{r}\left(w_h-\hat{E}(\bar{n}_{SIS.HH})\right)^2$$

R	is the number of (potential) initial transects within the survey region

$E(\bar{a}_{SIS.HH})$ Estimate for SIS Design

An estimate of the expected mean number of animals observed per transect leg is given by

$$\hat{E}(\bar{a}_{SIS.HH}) = \frac{1}{r}\sum_{h=1}^{r} u_h \tag{3.16}$$

where

$$u_h = \frac{1}{M}\sum_{i=1}^{K_h} \frac{\sum_{j \in \psi_i}\sum_{X=1}^{n_{ij}} a_{ijX}}{t_i}$$

a_{ijX} is the number of animals in the X[th] observation of the j[th] transect leg of the i[th] network

The variance is estimated by

$$\hat{V}\left(\hat{E}(\bar{a}_{SIS.HH})\right) = \frac{s_W^2}{r}\left(1 - \frac{r}{R}\right) \tag{3.17}$$

where

$$s_W^2 = \frac{1}{r-1}\sum_{h=1}^{r}\left(u_h - \hat{E}(\bar{a}_{SIS.HH})\right)^2$$

Horvitz-Thompson-based Estimators

$E(\bar{n}_{HT})$ Estimate for RIS Design

From equations 2.13 and 2.14, an estimate of the expected mean number of observations per transect for an initial sample without replacement is

$$\hat{E}(\bar{n}_{HT}) = \frac{1}{K}\sum_{i=1}^{v} \frac{\sum_{j=1}^{m_i} n_{ij}}{\alpha_i} \tag{3.18}$$

where

$$\alpha_i = 1 - \binom{K - m_i}{k} \Big/ \binom{K}{k}$$

K is the total number of transects in the survey region

v is the number of distinct networks in the sample

m_i is the number of transects in the i[th] network.

n_{ij} is the number of observations on the j[th] transect of the i[th] network

α_i is the probability that the ith network is included in the sample

k is the number of initial transects

If the initial sampling is without replacement, the variance is estimated by

$$\hat{V}\left(\hat{E}(\bar{n}_{HT})\right)=\frac{1}{K^2}\sum_{i=1}^{v}\sum_{h=1}^{v}\frac{\sum_{j=1}^{m_i}n_{ij}\sum_{j=1}^{m_h}n_{hj}\left(\alpha_{ih}-\alpha_i\alpha_h\right)}{\left(\alpha_i\alpha_h\alpha_{ih}\right)} \tag{3.19}$$

with

$$\alpha_{ii}=\alpha_i$$

and

$$\alpha_{ih}=1-\left\{\binom{K-m_i}{k}+\binom{K-m_h}{k}-\binom{K-m_i-m_h}{k}\right\}\Big/\binom{K}{k}$$

If the initial sampling is with replacement, then from Thompson (1996: p100)

$$\alpha_i=1-\left(1-\frac{m_i}{K}\right)^k$$

and

$$\alpha_{ih}=1-\left\{\left(1-\frac{m_i}{K}\right)^k+\left(1-\frac{m_h}{K}\right)^k-\left(1-\frac{m_i+m_h}{K}\right)^k\right\}$$

$E(\bar{a}_{HT})$ Estimate for RIS Design

An estimate of the expected mean number of animals observed per transect is given by

$$\hat{E}(\bar{a}_{HT})=\frac{1}{K}\sum_{i=1}^{v}\frac{\sum_{j=1}^{m_i}\sum_{X=1}^{n_{ij}}a_{ijX}}{\alpha_i} \tag{3.20}$$

where

a_{ijX} is the number of animals observed in the Xth observation of the jth transect of the ith network.

The variance is estimated by

$$\hat{V}\left(\hat{E}(\bar{a}_{HT})\right)=\frac{1}{K^2}\sum_{i=1}^{v}\sum_{h=1}^{v}\frac{\sum_{j=1}^{m_i}\sum_{X=1}^{n_{ij}}a_{ijX}\sum_{j=1}^{m_h}\sum_{X=1}^{n_{hj}}a_{hjX}\left(\alpha_{ih}-\alpha_i\alpha_h\right)}{\left(\alpha_i\alpha_h\alpha_{ih}\right)} \tag{3.21}$$

with α_i and α_{ih} as in Equations 3.18 and 3.19.

$E(\bar{n}_{SIS.HT})$ Estimate for SIS Design

From equations 2.21 and 2.22 an estimate of expected mean number of observations per transect leg is

$$\hat{E}(\bar{n}_{SIS.HT}) = \frac{1}{MR}\sum_{i=1}^{v}\frac{\sum_{j=1}^{m_i} n_{ij}}{\alpha_i} \tag{3.22}$$

where

$$\alpha_i = 1 - \binom{R-t_i}{r} \Big/ \binom{R}{r}$$

M	is the number of transects legs in each initial transect (primary unit)
R	is the number of initial transects in the survey region
v	is the number of distinct networks in the sample
m_i	is the number of transect legs in the ith network
n_{ij}	is the number of observations at the jth transect leg of the ith network
α_i	is the probability that the ith network is included in the sample
t_i	is the number of initial transects that intersect the ith network
r	is the number of initial transects in the sample

with variance

$$\hat{V}\left(\hat{E}(\bar{n}_{SIS.HT})\right) = \frac{1}{M^2R^2}\sum_{i=1}^{v}\sum_{h=1}^{v}\frac{\sum_{j=1}^{m_i}\sum_{j'=1}^{m_h} n_{ij}n_{ij'}(\alpha_{ih} - \alpha_i\alpha_h)}{(\alpha_i\alpha_h\alpha_{ih})} \tag{3.23}$$

where

$$\alpha_{ih} = 1 - \left\{\binom{R-t_i}{r} + \binom{R-t_h}{r} - \binom{R-t_i-t_h+t_{ih}}{r}\right\} \Big/ \binom{R}{r}$$

and

t_{ih}	is the number of initial transects that intersect networks i and h

$E(\bar{a}_{SIS.HT})$ **Estimate for SIS Design**

An estimate of the mean number of animals observed per transect leg is

$$\hat{E}(\bar{a}_{SIS.HT}) = \frac{1}{MR}\sum_{i=1}^{v}\frac{\sum_{j=1}^{m_i}\sum_{X=1}^{n_{ij}} a_{ijX}}{\alpha_i} \tag{3.24}$$

where

$$\alpha_i = 1 - \binom{R-t_i}{r} \bigg/ \binom{R}{r}$$

a_{ijX} is the number of observations in the X^{th} observation of the j^{th} transect leg of the i^{th} network

with variance

$$\hat{V}\left(\hat{E}(\bar{a}_{SIS.HT})\right) = \frac{1}{M^2R^2}\sum_{i=1}^{v}\sum_{h=1}^{v}\frac{\sum_{j=1}^{m_i}\sum_{X=1}^{n_{ij}}\sum_{j'=1}^{m_h}\sum_{X'=1}^{n_{ij'}} a_{ijX}\, a_{hj'X'}\left(\alpha_{ih} - \alpha_i\alpha_h\right)}{\left(\alpha_i\alpha_h\alpha_{ih}\right)} \tag{3.25}$$

3.2.5 Unequal Probability of Detection

The estimators so far defined in this chapter have assumed that the probability of detection is equal for all animals in the survey and that detection is purely dependent on an object's perpendicular distance from the trackline. Thus a single estimate of $f(0)$ has been formed by pooling observations across both initial and adaptive transects.

As already discussed, in many cases this will be unrepresentative and there are likely to be many other factors involved, such as group size, weather, observer experience, etc. So following the same process as for point transect sampling, we now produce an estimator which allows each observation to have its own probability of detection.

We first derive the probability p_Y, that the Y^{th} observation was detected within truncation distance W of the trackline, where the detection function is estimated using covariates (see, for example, Borchers 1996; Buckland *et al.*, 2002; Marques 2001). We then build on the Horvitz-Thompson-based estimators for a SIS design, used in equations 3.22 to 3.25, to define new estimators for the total number of animal groups in the surveyed area τ, and the total number of animals in the surveyed area δ. These are used to obtain estimates of both the animal group density and the individual animal density. Finally we also convert back to estimates for the animal group density and the individual animal densities.

Covariates and *f*(0).

As in Chapter 2 for point transects, let $\mathbf{z}_Y$ ($\mathbf{z}_Y = z_{i1}, z_{Y2}, \ldots, z_{Yq}$) be a matrix of the associated covariates for the Y^{th} observation so that its detection function is $g(x \mid \mathbf{z}_Y)$, at distance x from the line. We switch from indexing with X, as in the previous chapter, to Y, as we now use x to represent the distance from the line.

We first assume that objects are uniformly distributed, with respect to distance from the trackline, across the strip of truncation half-width, W. In fact this assumption is stronger than required, as shown by Fewster and Buckland (in preparation), and it is only required that the object distances have a linear probability distribution function.

Thus with truncation half-width W, the probability that an animal is at distance x or less from the line is

$$P(X \le x) = \frac{x}{W}, \qquad 0 \le x \le W$$

So the probability density function of the distance, x to each animal (whether detected or not) from the line is

$$u(x) = \frac{d\{P(X \le x)\}}{dx} = \frac{1}{W} \tag{3.26}$$

The probability p_Y, that object Y is detected within truncation width of W of the track line, and conditional on the observed values $\mathbf{z}_Y$, is given by

$$\begin{aligned} p_Y &= E_x\left[g(x \mid \mathbf{z}_Y) \mid \mathbf{z}_Y\right] \\ &= \int_0^W g(x \mid \mathbf{z}_Y) \cdot u(x \mid \mathbf{z}_Y)\, dx \end{aligned}$$

Assuming the x and $\mathbf{z}_Y$ are independent so that

$$p_Y = \int_0^W g(x \mid \mathbf{z}_Y) \cdot u(x)\, dx$$

and then substituting for $u(x)$ we get

$$\begin{aligned} p_Y &= \int_0^W g(x \mid \mathbf{z}_Y) \cdot \frac{1}{W}\, dx \\ &= \frac{1}{W} \int_0^W g(x \mid \mathbf{z}_Y)\, dx \end{aligned} \tag{3.27}$$

and an estimate of p_Y is given by

$$\hat{p}_Y = \frac{1}{W} \int_0^W \hat{g}(x \mid \mathbf{z}_Y)\, dx \tag{3.28}$$

where

$\hat{g}(x \mid \mathbf{z}_Y)$ is an estimate of the detection function

Defining the probability density function to be $f(x \mid \mathbf{z}_Y)$, then

$$\begin{aligned} f(x \mid \mathbf{z}_Y) dx &= \mathrm{pr}[\{\text{object in } (x, x+dx) \mid \mathbf{z}_Y\} \mid \{\text{object detected} \mid \mathbf{z}_Y\}] \\ &= \frac{\mathrm{pr}[\{\text{object in } (x, x+dx) \mid \mathbf{z}_Y\} \cap \{\text{object detected} \mid \mathbf{z}_Y\}]}{\mathrm{pr}\{\text{object detected} \mid \mathbf{z}_Y\}} \\ &= \frac{\mathrm{pr}[\{\text{object detected} \mid \mathbf{z}_Y\} \mid \{\text{object in } (x, x+dx) \mid \mathbf{z}_Y\}] \cdot \mathrm{pr}[\{\text{object in } (x, x+dx) \mid \mathbf{z}_Y\}]}{p_Y} \\ &= \frac{\mathrm{pr}[\{\text{object detected} \mid \mathbf{z}_Y\} \mid \{\text{object in } (x, x+dx) \mid \mathbf{z}_Y\}] \cdot \mathrm{pr}[\{\text{object in } (x, x+dx)\}]}{p_Y} \\ &= \frac{g(x \mid \mathbf{z}_Y) \cdot \left(\dfrac{dx \cdot L}{W \cdot L} \right)}{\dfrac{1}{W} \int_0^W g(x \mid \mathbf{z}_Y)\, dx} \\ &= \frac{g(x \mid \mathbf{z}_Y) \cdot dx}{\int_0^W g(x \mid \mathbf{z}_Y)\, dx} \end{aligned}$$

so that

$$f(x \mid \mathbf{z}_Y) = \frac{g(x \mid \mathbf{z}_Y)}{\int_0^W g(x \mid \mathbf{z}_Y)\, dx}$$

and thus

$$\int_0^W g(x \mid \mathbf{z}_Y)\, dx = \frac{g(x \mid \mathbf{z}_Y)}{f(x \mid \mathbf{z}_Y)}$$

So that if detection on the line is certain, and thus $g(0 \mid \mathbf{z}_Y) = 1$, then

$$\int_0^W g(x \mid \mathbf{z}_Y)\, dx = \frac{g(0 \mid \mathbf{z}_Y)}{f(0 \mid \mathbf{z}_Y)} = \frac{1}{f(0 \mid \mathbf{z}_Y)} \tag{3.29}$$

Finally, substituting 3.29 in to 3.28, an estimate of p_Y, if detection on the trackline is certain, is given by

$$\hat{p}_Y = \frac{1}{W \cdot \hat{f}(0 \mid \mathbf{z}_Y)} \tag{3.30}$$

where

$\hat{f}(0 \mid \mathbf{z}_Y)$ is an estimate of the probability density function of detection distances for the Y^{th} observation with covariates $\mathbf{z}_Y$, evaluated at zero

SIS Design: Horvitz-Thompson-based Estimate for Unequal Probability of Detection

We want to define a Horvitz-Thompson-based estimator for a SIS design, where each animal has a unique probability of detection. Thus using the estimator of equation 1.8 as a basis, and following the approach of Thompson and Seber (1996: p226), an estimate of number of animals in the surveyed areas given perfect detectability is

$$\hat{\delta}_0 = \sum_{i=1}^{V} \frac{s_i^*}{\alpha_i}$$

where

$$s_i^* = \sum_{j=1}^{m_i} \sum_{Y=1}^{n_{ij}} s_{ijY}$$

s_{ijY} is the number of animals in the Y^{th} animal group detected on the j^{th} transect leg of the i^{th} network

Letting $\hat{u}_{ijY}$ denote the value of a new variable, using estimated detectability, so that

$$\hat{u}_{ijY} = \frac{s_{ijY} I_{ijY}}{\hat{p}_{ijY}}$$

where

I_{ijY} is an indicator variable, such that I_{ijY} is 1 if the Y^{th} animal group of the j^{th} transect leg of the i^{th} network is detected and 0 otherwise

$\hat{p}_{ijY}$ is an estimate of the probability of detection of the Y^{th} object of the j^{th} transect leg of the i^{th} network

Then by Thompson and Seber (1996: p226) the total number of individual animals in the surveyed area, $\hat{\delta}_{UP.SIS.HT}$ can be estimated by

$$\begin{aligned}\hat{\delta}_{UP.SIS.HT} &= \hat{\delta}_0(\hat{\mathbf{u}}_s) \\ &= \sum_{i=1}^{v} \frac{\hat{u}_i^*}{\alpha_i}\end{aligned} \tag{3.31}$$

where

$$\hat{\mathbf{u}}_s = \left\{ \hat{u}_{ijY} : i \in s \right\}$$

$$\hat{u}_i^* = \sum_{j=1}^{m_i} \sum_{Y=1}^{n_{ij}} \frac{s_{ijY}}{\hat{p}_{ijY}}$$

$$\alpha_i = 1 - \binom{R - t_i}{r} \Big/ \binom{R}{r}$$

v is the number of distinct networks in the sample

m_i is the number of units in the i^{th} network

n_{ij} is the number of animal groups detected on the j^{th} transect leg in the i^{th} network

s_{ijY} is the number of animals in the Y^{th} animal group detected on the j^{th} transect leg of the i^{th} network

$\hat{p}_{ijY}$ is an estimate of the probability of detection for the Y^{th} animal group on the j^{th} transect leg of the i^{th} network

Thus by Thompson and Seber (1996: p227), an estimate of the variance of $\hat{\delta}_{UP.SIS.HT}$ is given by

$$\hat{V}\left(\hat{\delta}_{UP.SIS.HT}\right) = \hat{v}_1 + \hat{v}_2 + \hat{v}_3$$

with

$$\begin{aligned}\hat{v}_1 &= v_0(\hat{\mathbf{u}}_s) \\ &= \sum_{i=1}^{v} \sum_{i'=1}^{v} \hat{u}_i^* \hat{u}_{i'}^* \frac{(\alpha_{ii'} - \alpha_i \alpha_{i'})}{\alpha_{ii'} \alpha_i \alpha_{i'}}\end{aligned} \tag{3.32}$$

where

$$\alpha_{ii'} = 1 - \left\{ \binom{R - t_i}{r} + \binom{R - t_{i'}}{r} - \binom{R - t_i - t_{i'} + t_{ii'}}{r} \right\} \Big/ \binom{R}{r}$$

and if

$$\hat{\beta}_{ijY} = \frac{(1 - \hat{p}_{ijY})}{\hat{p}_{ijY}^2} I_{ijX} s_{ijY}^2$$

then

$$\begin{aligned} \hat{v}_2 &= \hat{\tau}_0(\hat{\boldsymbol{\beta}}_s) \\ &= \sum_{i=1}^{v} \sum_{j=1}^{m_i} \sum_{Y=1}^{n_{ij}} \frac{(1 - \hat{p}_{ijY})}{\alpha_i \cdot \hat{p}_{ijY}^2} s_{ijY}^2 \end{aligned} \tag{3.33}$$

and finally

$$\hat{v}_3 = \sum_{i=1}^{v} \sum_{i'=1}^{v} \sum_{j=1}^{m_i} \sum_{j'=1}^{m_{i'}} \sum_{Y=1}^{n_{ij}} \sum_{Y'=1}^{n_{i'j'}} \frac{s_{ijY} s_{i'j'Y'}}{\alpha_{ii'} \hat{p}_{ijY}^2 \hat{p}_{i'j'Y'}^2} \hat{\text{cov}}(\hat{p}_{ijY}, \hat{p}_{i'j'Y'}) \tag{3.34}$$

Thus

$$\begin{aligned} \hat{V}(\hat{\delta}_{UP.SIS.HT}) = &\sum_{i=1}^{v} \sum_{i'=1}^{v} \hat{u}_i^* \hat{u}_{i'}^* \frac{(\alpha_{ii'} - \alpha_i \alpha_{i'})}{\alpha_{ii'} \alpha_i \alpha_{i'}} \\ &+ \sum_{i=1}^{v} \sum_{j=1}^{m_i} \sum_{Y=1}^{n_{ij}} \frac{(1 - \hat{p}_{ijY})}{\alpha_i \cdot \hat{p}_{ijY}^2} s_{ijY}^2 \\ &+ \sum_{i=1}^{v} \sum_{i'=1}^{v} \sum_{j=1}^{m_i} \sum_{j'=1}^{m_{i'}} \sum_{Y=1}^{n_{ij}} \sum_{Y'=1}^{n_{i'j'}} \frac{s_{ijY} s_{i'j'Y'}}{\alpha_{ii'} \hat{p}_{ijY}^2 \hat{p}_{i'j'Y'}^2} \hat{\text{cov}}(\hat{p}_{ijY}, \hat{p}_{i'j'Y'}) \end{aligned} \tag{3.35}$$

Substituting equation 3.28 for the $\hat{p}_{ijY}$ in 3.35 gives

$$\hat{\delta}_{UP.SIS.HT} = \sum_{i=1}^{v} \frac{\hat{u}_i^*}{\alpha_i} \tag{3.36}$$

where

$$\hat{u}_i^* = W \cdot \sum_{j=1}^{m_i} \sum_{Y=1}^{n_{ij}} \left(\frac{s_{ijY}}{\int_0^W \hat{g}(x \mid \mathbf{z}_{ijY})\, dx} \right)$$

W is the truncation half-width

If the probability of detection on the trackline is certain, $g(0 \mid \mathbf{z}_{ijY}) = 1$, then from equation 3.30 this can be simplified to

$$\hat{\delta}_{UP.SIS.HT} = W \cdot \sum_{i=1}^{v} \left(\sum_{j=1}^{m_i} \sum_{Y=1}^{n_{ij}} s_{ijY} \cdot \hat{f}(0 \mid \mathbf{z}_{ijY}) \Big/ \alpha_i \right) \tag{3.37}$$

An estimate of the total number of animal groups in the surveyed area can be formed by replacing the group size, s_{ijY} by 1, so that

$$\hat{\tau}_{UP.SIS.HT} = \sum_{i=1}^{v} \frac{\tilde{u}_i^*}{\alpha_i} \tag{3.38}$$

where

$$\tilde{u}_i^* = \sum_{j=1}^{m_i} \sum_{X=1}^{n_{ij}} \frac{1}{\hat{p}_{ijY}}$$

and

$$\begin{aligned} \hat{V}(\hat{\tau}_{UP.SIS.HT}) = & \sum_{i=1}^{v} \sum_{i'=1}^{v} \tilde{u}_i^* \tilde{u}_{i'}^* \frac{(\alpha_{ii'} - \alpha_i \alpha_{i'})}{\alpha_{ii'} \alpha_i \alpha_{i'}} \\ & + \sum_{i=1}^{v} \sum_{j=1}^{m_i} \sum_{Y=1}^{n_{ij}} \frac{(1 - \hat{p}_{ijY})}{\alpha_i \cdot \hat{p}_{ijY}^2} \\ & + \sum_{i=1}^{v} \sum_{i'=1}^{v} \sum_{j=1}^{m_i} \sum_{j'=1}^{m_{i'}} \sum_{Y=1}^{n_{ij}} \sum_{Y'=1}^{n_{i'j'}} \frac{1}{\alpha_{ii'} \hat{p}_{ijY}^2 \hat{p}_{i'j'Y'}^2} \hat{\text{cov}}(\hat{p}_{ijY}, \hat{p}_{i'j'Y'}) \end{aligned} \tag{3.39}$$

As with the $\hat{\delta}_{UP.SIS.HT}$ estimate, this can be further simplified if $g(0 \mid \mathbf{z}_{ijY}) = 1$, by replacing the estimate of p_{ijY} by its simplified form in equation 3.30.

Density Estimates

There are K transect legs in the survey region each with length l and truncation half-width W, so by equation 3.3, the surveyed area is given by $2lWK$. As $\hat{\tau}_{UP.SIS.HT}$ is an estimate of the total number of groups in the surveyed area, then an estimate of the group density is given by

$$\hat{D}_{UP.g} = \frac{\hat{\tau}_{UP.SIS.HT}}{2lWK} \tag{3.40}$$

where

K is the number of transects legs in the survey area

with variance estimated by

$$\hat{V}(\hat{D}_{UP.g}) = \frac{\hat{V}(\hat{\tau}_{UP.SIS.HT})}{4l^2 W^2 K^2} \tag{3.41}$$

The individual animal density estimate is

$$\hat{D}_{UP} = \frac{\hat{\delta}_{UP.SIS.HT}}{2lWK} \tag{3.42}$$

and an estimate of variance is given by

$$\hat{V}\left(\hat{D}_{UP}\right) = \frac{\hat{V}\left(\hat{\delta}_{UP.SIS.HT}\right)}{4l^2W^2K^2} \tag{3.43}$$

As the variance estimate above makes many assumptions about the independence of variables and is likely to underestimate the variance, in practice a bootstrap based estimate will be preferred. This is discussed later in the chapter.

Abundance Estimates

Defining A as the area of the survey region and A_S as the total survey area, then from equation 2.4, an estimate of the total number of individual animals in the survey region is given by

$$\hat{N}_{UP.SIS.HT} = \frac{A}{A_S} \cdot \hat{\delta}_{UP.SIS.HT} \tag{3.44}$$

and its variance is estimated by

$$\hat{V}\left(\hat{N}\right) = \hat{V}\left(\frac{A}{A_S} \cdot \hat{\delta}_{UP.SIS.HT}\right) = \frac{A^2}{A_S^2} \cdot \hat{V}\left(\hat{\delta}_{UP.SIS.HT}\right) \tag{3.45}$$

Similarly, from equation 2.3, an estimate of the total number of animal groups in the survey region is given by

$$\hat{N}_g = \frac{A}{A_S} \cdot \hat{\tau}_{UP.SIS.HT} \tag{3.46}$$

and its variance is estimated by

$$\hat{V}\left(\hat{N}_g\right) = \frac{A^2}{A_S^2} \cdot \hat{V}\left(\hat{\tau}_{UP.SIS.HT}\right) \tag{3.47}$$

3.3 Grid and Neighbourhood Design

There are fewer options to consider for the grid design than for the adaptive point transects in Chapter 2. The sampling strip is rectangular, so as already stated, the obvious choice for the grid units is also a rectangle. Thus although other shapes are possible, we do not consider them here.

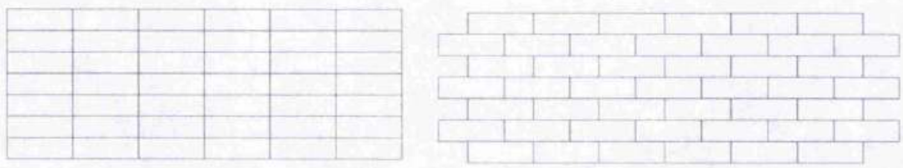


Figure 3.10: Rectangular units assembled to create a grid, either aligned as in the diagram on the left or offset as in the diagram on the right.

The rectangular units can either be assembled as a conventional grid or offset as shown in Figure 3.10. The offset grid allows some new variants of adaptive pattern to be considered, such as adding the two adjacent rectangles on either side of a triggering unit. However although there may be some benefits where a specific spatial distribution is being encountered, for the majority of cases the standard grid will suffice. Hence we now consider adaptive patterns using a conventional grid of rectangles.

3.3.1 Adaptive Patterns

The adaptive (neighbourhood) pattern will be dependent on the scale of the transects or transect legs and the overall survey design. So, for example, if the survey has very long transects or transect legs, in comparison to the total survey length, then it will be sensible to keep the number of adaptive units in a neighbourhood small. Conversely, if the transects or transect legs are short then it may be feasible to use more transects in the neighbourhood pattern. A neighbourhood must be made up of complete units. So if, in the terminology of this chapter, the survey is a RIS design, each unit of the grid represents a transect, whilst if it is a SIS design, each unit represents a transect leg. Thus for RIS designs, transects are added to form the neighbourhood and for SIS designs, transect legs are used.

Although there are many variations, four basic adaptive patterns form the basis of most of these. The adaptive pattern should be such that any unit in a network should also include all other units within the network, and thus needs to be symmetrical in at least one direction. The four patterns are referred to as the *NSEW* (North, South, East and West), *Parallel*, *X-Shape* and *All Units* neighbourhoods.

NSEW Neighbourhood

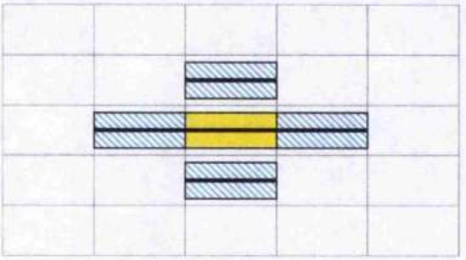


Figure 3.11: Example NSEW neighbourhood pattern using a rectangular grid. Tracklines are shown as a black line, the initial transects are signified with a solid (yellow) sampling strip, and the adaptive units by a (blue) cross-hatched sampling strip.

This adaptive pattern, shown in Figure 3.11, is typically used in examples of Thompson's methods and adds the units above, below and to either side of the triggering unit (the north, south east and west of the name).

A major factor with line transect surveys is the amount of off-effort travel the pattern involves. For many line transect surveys the rectangles are likely to be long and thin, whilst for ease of representation they have been shown as relatively short. Thus the NSEW neighbourhood is likely to introduce a large amount of off-effort travel as the observer moves between the adaptive transects. It also adds four times the effort to the triggering transect or transect leg and so care will be need to be taken in selecting the trigger condition.

The pattern will typically be used with non-contiguous transects or transects legs, and does have the advantage that it allows the adaptive transects or legs to follow a cluster both perpendicular to and in the direction of the trackline. However it could be used with contiguous transect legs in a SIS design to allow the adaptive transect legs to follow a spatial cluster of animals in more than one direction. In this case with a triggering transect leg in the middle of a transect, the first adaptive legs will only be added to the parallel units on either side of the initial leg, as the legs in front and behind are initial legs and so are already included in the survey. However if either of the adaptive legs were to trigger, then these would also add additional legs on the three un-surveyed edges of the unit and so provide the opportunity to detect objects in multiple directions.

This pattern is likely to be of benefit where there are dense but sparsely located clusters so that on encountering a cluster there will be benefit in an extensive survey of the surrounding area.

Parallel Neighbourhood

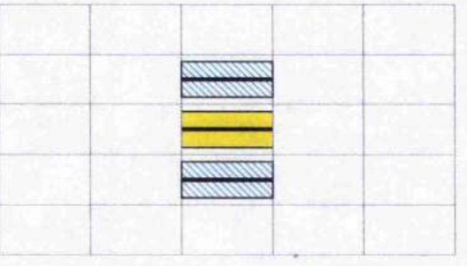


Figure 3.12: Example Parallel neighbourhood pattern using a rectangular grid.

This pattern, as the name suggests, adds the units on either side of the triggering unit, so that the adaptive transects or transect legs run parallel with the triggering trackline (Figure 3.12). The main advantage is it only adds the minimum number of adaptive units; two on the trigger from an initial transect or transect leg, and one for each adaptive unit that also triggers. Thus it will typically be used where there are a low number of initial transects or legs, to keep the proportion of adaptive effort down. It also uses significantly less off-effort than the NSEW neighbourhood - for the first trigger of an initial transect or leg, it is possible to minimise the off-effort to just the perpendicular movement between transects. This is particularly useful if used for contiguous transect legs in a SIS design. In this situation if the trigger condition is met at the end of an initial transect leg, the observer moves, perpendicular to the transect leg, to the adaptive transect leg on one side and surveys this in the opposite direction. The observer then crosses over the initial transect leg to the adaptive transect leg on the other side. This is surveyed in the original direction, so that at the end of it the observer moves back to the initial trackline and is in the correct position to survey the next initial transect leg (Figure 3.13). The movement between legs becomes more complex if the adaptive legs also trigger adaptive effort. The order in which legs are surveyed is discussed further in section 6.2, Survey Design Considerations.

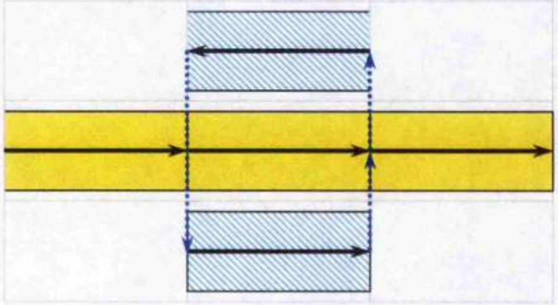


Figure 3.13: Movement between transect legs with a Parallel neighbourhood. The arrow head signifies the surveying direction on the transect legs. The blue dashed lines represent the off-effort travel.

A downside of the pattern is that it will only follow a spatial cluster of animals to either side of the trackline, so that if the initial transect or transect leg just clipped the edge of a cluster, it is less likely the adaptive units will cover a large amount of the cluster.

X-Shape Neighbourhood

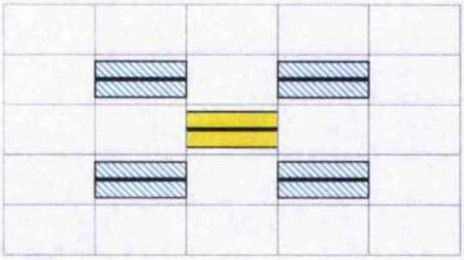


Figure 3.14: Example X-Shape neighbourhood pattern using a rectangular grid.

This pattern adds the units at the corners of the triggering unit, so that they form an X, as shown in Figure 3.14. As with the NSEW neighbourhood it adds four times the effort to the initial triggering transect or transect leg and can follow a spatial cluster of animals in multiple directions. It has two main differences from the NSEW neighbourhood. It has gaps between surveyed units, in a checkerboard effect, and as a result will spread more quickly over an area. Thus it will fair better where the spatial cluster of animals is less dense and spread over a fairly large area compared to the size of a transect or transect leg. Otherwise with dense clusters, the pattern risks stepping over and thus outside of a cluster. The other difference is that if used with a SIS design of contiguous transect legs, then a triggered initial transect adds four adaptive transect legs, whilst for the same design an NSEW neighbourhood only adds two adaptive legs to an initial leg.

All Units Neighbourhood

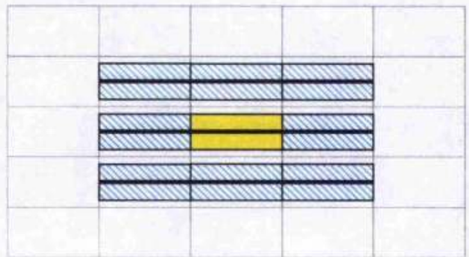


Figure 3.15: Example All Units neighbourhood pattern using a rectangular grid.

The last pattern considered is a combination of the X-Shape and NSEW neighbourhoods where all the immediately surrounding units of the triggering unit are added (Figure 3.15). This adds a significant amount of adaptive effort, but has the ability to follow a spatial cluster of animals in any direction. It also uses a lower proportion of off-effort travel than either the X-Shape or the NSEW as it is possible to navigate it in a similar (but not identical) fashion to the Parallel neighbourhood. This neighbourhood would only be used where there was a large number of initial transects or transect legs and there were few, but dense, spatial clusters of the animals being surveyed.

3.3.2 Grid Design Selection and Application

As for point transects each survey will need to be carefully designed with reference to its particular characteristics. Many of the points relating to point transect grid design in section 2.3.5 also apply with line transects, and the section should be read in conjunction with this.

Unless there is specific reason otherwise, a standard rectangular grid should be used. If there is limited information on the surveying characteristics of the population, then the Parallel neighbourhood, which has the best all-round features, should be used. The low number of adaptive units added will be an advantage for multi-species surveys as, assuming triggering is on a single rare species, it will be preferable not to penalise the surveying of other species with excessive adaptive effort. The Parallel neighbourhood is also most appropriate when only a few long transects are used. The

NSEW and X-Shape neighbourhoods are suited to surveys consisting of many transects or transect legs and where off-effort does not use significant survey time or resource to move between the adaptive transects. The All Units neighbourhood should only be used when tracking an extremely rare species with high spatial clustering and additional adaptive effort is less of a consideration.

Although the methods allow for the truncation half-width to be such that the sampling strips in adjacent units overlap, this should be avoided if possible, as it will introduce complications to the field methods and may increase responsive movement.

For some surveys that cover large areas with slow moving observers, efficient use of survey resources is achieved by connecting a number of transects in a zigzag pattern and considering each transect as independent. Thus the tracklines do not all lie in the same direction and the survey area cannot be overlaid with a single grid. Assuming the transects are very long with a narrow strip width, then it may be acceptable to align a smaller grid with each transect. This is effectively dividing the survey area up into a number of smaller areas so that each transect falls within it own area and there is a separate grid for each area. Ideally the survey should be a SIS design, so that only short transect legs are added alongside the main transects. There will be a small edge effect at the ends of the transects, where the grids overlap, but if this is kept small then it should be acceptable. The units in each grid should be the same size, and the total number of units in the survey region should be considered as if the area was covered by a single grid.

3.4 Limiting Total Effort

A major limitation of Thompson's adaptive methods is that the total effort required is unknown at the start of the survey, and is a function of the value of the variable of interest in each sampling unit. Thus it can be difficult to forecast how long to book resources, such as observers or a survey ship. An incorrect choice of trigger could cause major problems for a survey. If the trigger is too high, then the survey may not adapt at all. This is probably less of an issue, as the sampling units would all become networks of a single unit, and the results would revert to conventional style estimates. If resources were being reserved for a fixed period, then it is likely that some additional time would have been factored in to account for the adaptive effort and thus the use of resources would not have been maximised. At the other extreme if the trigger is set too low, then the survey could potentially trigger on every sampling unit and attempt to cover the whole survey region. This will be a more serious issue as it is unlikely sufficient resources will be available, leading to uneven coverage of the survey region. In Figure 3.16, a trigger of 1 has added 19 adaptive transect legs to 6 nominal legs, with the number of detections increased from 14 to 33 by the adaptive procedure. However, increasing the trigger to 2 only adds 10 adaptive legs and yet still gives 31 detections in total. Thus the cost (in time and resources) of making the additional detections is greatly reduced.

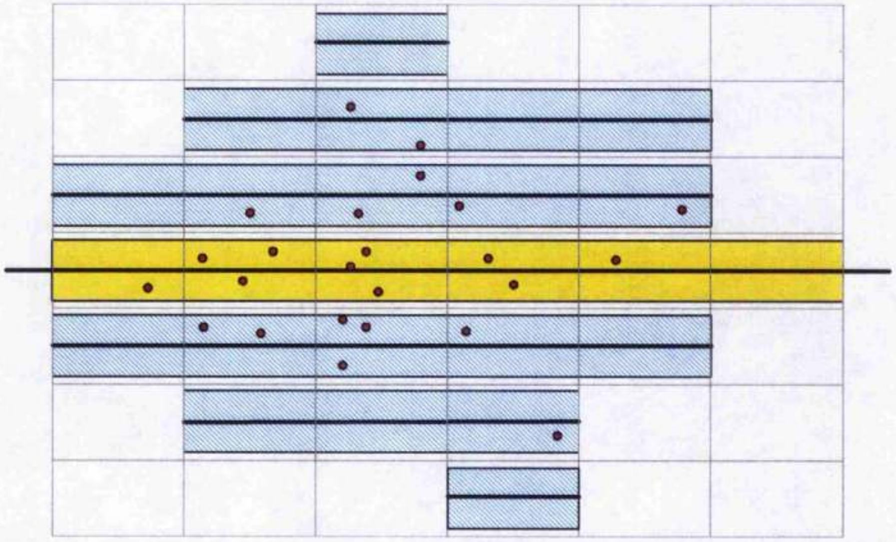


Figure 3.16: An incorrect trigger value can lead to excessive adaptive effort. Here a trigger value of 1, with a Parallel adaptive neighbourhood, has added 19 adaptive transect legs (shown as blue cross-hatched strips) and 14 detections (shown as red dots) to the initial 6 nominal legs (shown as a yellow strips) with 10 detections. Increasing the trigger to 2 would have only added 10 adaptive legs, but would still have generated 12 new detections (only a reduction of 2 with greatly reduced adaptive effort).

Ideally some previous survey data will be available to act as a guide. A simplistic approach would be to examine previous encounter rates and use this as the basis for selecting a trigger. Preferably the data can be used as the basis for simulations to verify the survey parameters are correct. If previous survey data are not available, then it may be possible to use biological knowledge of the species, or to arrange a short pilot survey.

Thompson and Seber (1996, p160) suggest approaches to limit the total sampling effort with their adaptive estimators, and these are discussed further here in relation to adaptive distance sampling.

One approach is to pre-stratify the survey region into a number of distinct areas. Thompson has shown that for each stratum the trigger condition can be set based upon information from the previous strata, such as effort expended, abundance observed or the variance estimate. Then for each stratum, design unbiased estimators of the stratum population and estimation variance are made according to the estimation approach used for the stratum. A design unbiased total population estimate is then given by the sum of the individual stratum estimates and the estimate of variance is the sum of the individual stratum variance estimates.

Thompson and Seber also describe a pragmatic approach where post-stratification is used during analysis. As already mentioned, the initially selected sampling units can be used to provide an unbiased estimate of the population, as if no adaptive sampling had been planned. Therefore it is sensible to aim to give priority to completion of the initial sample. Due to the geographical spread of the survey, or some other such factor, many surveys will be forced to complete the adaptive sections as they are encountered. In this case the minimum effort required to complete the remainder of the survey should always be checked prior to adapting. If at any point in time the remaining effort available falls to this minimum level, then the survey should be completed surveying only the remaining initial sampling units. The survey is then stratified into two strata: an adaptive stratum and a conventional stratum. Results are analysed with appropriate adaptive and conventional estimators, before combining as stratified estimates. If it was assumed that the adaptive process did not introduce any heterogeneity into the detection function estimation, then the sightings from both strata could be pooled to create a single $f(0)$ estimate. The approach can be extended to apply to changes in the trigger condition during the survey, so that if, for example, the trigger was too low and the survey was using too much adaptive effort, the trigger could be increased. The two (or more) areas would be estimated using the adaptive estimators and then combined as stratified estimates, as suggested for the pre-survey stratification. However in this case, the estimator will be biased. In such a situation it would be better to increase the trigger whilst there is still additional effort available as discussed in the following paragraph.

If the effort used in adapting is carefully monitored during the survey, it may be feasible to revise the trigger without a large impact on the survey. Thus if the trigger is too low, and so the survey adapts too much, then if identified early enough, it may be best to increase the trigger and drop any previously sampled units that do not meet this new condition, with little effort lost. Conversely if the trigger was too high and so has not adapted enough, it may be possible to set a lower trigger and revisit sampling units which would have been included with this revised trigger. This second case is more complex, and may not be viable with some surveys. Thus, for example, it may not be feasible to return to the area due to the off-effort involved; or the population may be transitory and the time window for surveying the area had passed.

Thompson and Seber also propose partitioning the region into blocks or regions, so that an adaptive section cannot step outside a block. This could be considered similar to sampling apartments in a city divided into blocks: apartments are added in an adaptive way within the block, but the procedure cannot step outside the block. A similar result could be achieved with line transect sampling by dividing the grid overlaying the survey region into blocks of transects, say five transects wide, and running any initial transect down the centre transect of the five. The initial transects are located systematically within the survey region, and are assumed to be independent as the grid is randomly located. Thus with a Parallel neighbourhood, a unit could trigger a maximum of twice in either direction before being stopped by the edge of the block. The same approach could be extended to have the blocks restrict access in two directions, to allow for more complex neighbourhoods, such as the NSEW neighbourhood. The approach is equally applicable to point transect sampling.

Finally, neighbourhoods like the X-Shape can be used, which add units in a checkerboard style, and thus cover a wider area, so that the total effort required to sample the area covered by a cluster is reduced. Thus even in the worst case of the trigger being too low and continually adapting, the total effort is kept within a manageable limit. This tactic will be dependent on the size of clusters expected in the population, and should only be considered as a safeguard rather than a direct method for limiting effort.

Other approaches such as the two-stage cluster sampling of Salehi and Seber (1997) are also possible, although these will involve different adaptive estimators.

3.5 Discussion

It is expected that these methods may have less benefit for line transects than for point transects, as for line transects, the density function estimate typically makes a smaller contribution to the overall density variance. As with point transects, it is not thought that the methods will introduce any significant field issues, although the additional time off-effort is likely to be a greater issue with line transects than for points.

Where movement between transects or transect legs is fast or low cost (either monetary or in survey resources), and surveying is comparatively slow and painstaking, then the overhead of the adaptive process is reduced. An example could be the deer dung pellet surveys described by Marques *et al.* (2001). Here short transects (transect legs in the terminology of this chapter) were performed in a systematic grid. Two basic survey designs were used. In one area transect legs were 200 metres long with gaps or either 200 or 400 metres between the end of one leg and the start of the next, dependent on the expected density in the area. The other design used 50 metre transect legs with gaps of 400, 600 or 800 metres between legs. Thus the survey can be considered a SIS design. The trackline was marked with a rope or cable with length marks and the observer slowly walked along it searching for deer dung pellets. The sampling strips were long and narrow, and the effective strip half-widths were typically between 1 and 2 metres. Counts on transects were often low as surveying was performed across a wide area, some of which contained no deer. Thus the survey has many characteristics that would make it suitable for an adaptive line transect survey using a NSEW neighbourhood. The justifications for this include: many transects experiencing a low encounter rate; observers could easily move to a neighbouring unit; responsive movement is not an issue; and within reason, adaptive transects can be performed in any order.

The Thompson-based adaptive line transect approach shares many of the issues faced for point transects, including: the complex notation and unwieldy estimation formulae; the potential to disturb animals and so experience bias from responsive behaviour; issues of selecting a suitable trigger; and the complications in handling multi-species surveys; particularly if species have differing spatial distributions.

As with the point transect estimators, it is possible to use bootstrap methods (Efron and Tibshirani, 1993) to estimate variances and confidence intervals. For a RIS design the sampling units will be the initial transects plus any adaptive units they include. If the adaptive units form a network, then the complete network and any edge units should be added. Although not used in the estimates of the mean number of animals and animal groups per transect or transect leg, observations made on edge units will need to be included for the detection function estimation. For a SIS design, the sampling units will again be the initial transects (the primary units), and when a primary unit is selected, then all related adaptive transect legs including the networks they form and the associated edge units should also be selected.

Care will need to be taken with the trigger condition as when units are continually added to a neighbourhood and the collection of adaptive units grows, so the off-effort may become increasingly cumbersome, with more complex routes for the observers to traverse between neighbourhoods. This will also increase the risk of animal disturbance. In addition, if the animals are mobile and the surveying takes too long, then there is a risk that the cluster of animals may have moved out of the survey area before sampling is completed, and the adapting may have been to no avail. Thus to minimise these effects the number of adaptive units should be kept low.

In the next chapter, we consider a new estimator that allows adaptive sampling to be performed using a fixed total effort.

Chapter 4

Adaptive Distance Sampling with Fixed Effort

4.1 Introduction

In Chapter 2 and Chapter 3 we combined Thompson's adaptive methods with point and line transect sampling. Although this approach improved precision, the total effort required cannot be reliably predicted in advance. So the planned survey time may expire long before the study area has been fully surveyed if more animals than expected are encountered; or available time is not fully utilized if fewer animals than expected are found.

For many surveys it will be preferable to complete the survey within a planned, fixed, total effort, so that resources and expenses can be identified prior to commencing the survey. In particular this will be useful where sampling uses expensive resources, such as the survey ship, observers and crew for a marine mammal survey. Also the survey platforms may only be available for a fixed time, with survey schedules open to disruption from external factors such as the weather, resulting in poor or uneven coverage of the survey area.

We now develop an adaptive line transect sampling approach with fixed total effort, which can both improve precision, for clustered populations, and also improve survey coverage. We refer to this approach as the PB method after the original authors, Pollard and Buckland (1997).

Thompson's methods typically use networks and neighbourhoods (Chapter 1) which can be difficult to accommodate in distance sampling, and introduce a degree of *off-effort*, whilst the observer moves between units in a neighbourhood. The approach here uses less complex notation and depending on the adaptive pattern selected, can allow the adaptive sections to follow immediately from a conventional section, with minimal or no off-effort. This further increases the survey efficiency.

The basic approach is to increase survey effort in areas of high abundance, based on a trigger condition (for example the encounter rate rising above some value), as with Thompson's methods. With this approach the increase in sampling intensity is a function of the degree to which the observer is ahead of or behind schedule. This increase in sampling intensity is measured by the *effort factor*. So, if effort is doubled when the trigger is activated, the effort factor is 2 for that section of the survey. The effort factor is then used to weight sightings to remove bias.

The approach conditions on the effort factors, which are data-dependent, and thus the method is not design-unbiased (Thompson, 1992: p17). However simulations have shown that negligible bias is introduced and that the methods represent an acceptable procedure.

Although many adaptive patterns are feasible, for this chapter we concentrate on the use of a zigzag track (Figure 4.1), as this requires no additional off-effort and is easily implemented.

We start by developing estimation formulae for adaptive line transect sampling. The formulae are derived for passing-mode surveys only, where the observation platform (e.g. ship) does not detour to investigate observations. It may be possible to modify the formulae to accommodate closing-mode surveys, but this has not been investigated by the author. The section is primarily written from the perspective of a marine line transect survey, but the strategy is considered potentially applicable to all platforms, marine, airborne and terrestrial.

Following the course of the earlier chapters, we initially pool observations to estimate $f(0)$, and then proceed to develop an $f(0)$ estimator that includes covariates. The theory is tested through simulation and later, in Chapter 5, the methods are applied to an experimental harbour porpoise survey to assess both the theory and practicality of the approach.

Although the approach is developed for line transect sampling, and in particular shipboard surveys, the chapter also extends the method for point transects, before considering extensions and closing with a discussion of the methods.

Within this chapter, as in the rest of the thesis, group or school size is used to refer to the number of animals in a single detected group, while cluster refers to a spatial cluster of animal groups (where a 'group' may comprise just one animal).

4.2 Line Transect Theory

The underlying principle of adaptive line transect sampling is that the sampling effort is increased in areas of high animal density, leading to larger sample sizes and thus improved estimator precision. Furthermore, in this implementation of adaptive sampling, the increase in effort is a function of the available effort remaining, improving coverage for surveys that have to complete within a fixed timescale.

A minimum amount of survey effort, termed the *nominal effort*, is predetermined. This is the effort required, without any adapting, to complete the survey as a conventional survey. The total effort is set based on the nominal effort plus an amount of additional effort for the adaptive sections. At any time within the survey the degree to which survey effort increases in areas of high density is a function of the difference between the total effort still available and the nominal effort remaining.

The survey effort is adapted by increasing the effort, above the nominal straight line effort, when the number of observations exceeds some limit. The increased adaptive effort is maintained for some period, after which the observer returns to the nominal (straight line) track. The amount of time or nominal distance for which the increased effort is maintained is a parameter of the survey and will affect the survey efficiency; this is discussed later in the chapter.

4.2.1 Adapting the Nominal Effort

The increase in effort is measured by the *effort factor*, represented by λ. The effort factor is defined to be the ratio of the effort used following an adaptive track, relative to the effort that would be used to follow the corresponding straight-line (nominal) track. Thus a transect may be divided into a number of sub-transects, or *legs*, each with a different effort factor. The formulae derived in this chapter are conditional on the effort factors, and are therefore not *design-unbiased* (Thompson, 1992: p17). However, later in the chapter, simulation results demonstrate that little bias is introduced by this conditioning.

Using conventional line transect estimators, systematically increasing the effort in areas of higher animal density would lead to abundance overestimation. The method avoids this difficulty by downweighting data from adaptive sections, to compensate for increased effort. The weight is inversely proportional to the effort factor, so that each section of transect is given weight in proportion to the length of straight-line (*nominal*) effort through that section. We define the nominal number of observations in a section to be the number of detections had the nominal search effort been carried out. This is estimated in adaptive sections by dividing the actual number by the effort factor.

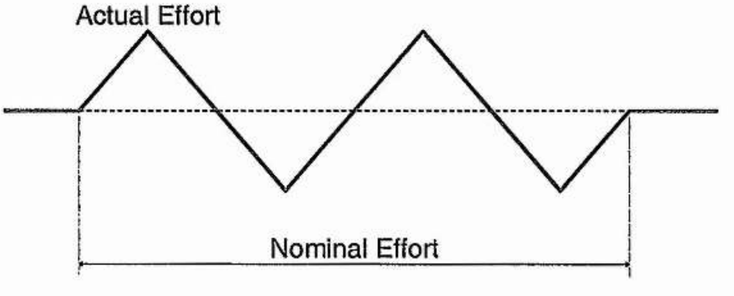


Figure 4.1: Increasing effort by zigzagging, nominal effort refers to the equivalent straight line track.

Many tracking designs for increasing line transect effort are possible (e.g. zigzag, hounds-tooth, sinusoidal). Here we concentrate on the zigzag pattern (Figure 4.1) as it has a number of advantages. In particular, the trackline does not cross itself; there are no gaps in the trackline so that no search effort is lost in travelling from one transect to the next; and the track is easily followed (important for shipboard surveys). In addition, the increase in effort is directly related to the length, angle and number of the zigzags and thus can be fixed at any value ≥ 1 by changing these factors, either singularly or in combination.

In conventional non-adaptive designs, poor weather conditions may lead to gaps in the nominal effort, or a failure to cover the entire study area. With Thompson's methods, the increase in effort is predetermined, whilst the number of times increased effort is triggered (the survey adapts) is a random variable. Thus in line transect surveys, the entire area might be completed with time to spare, reducing efficiency, or worse, time may run out before the area has been fully covered, leading to unrepresentative cover and bias. In this method, a rule is used to ensure that the 'effort factor' is a function of how much the survey is ahead or behind schedule.

4.2.2 Notation

Each transect is divided into a number of sub-transects or legs, where the start and finish of each leg occurs at a change in effort, as shown in Figure 4.2 for a zigzag track.

L	is used to represent the total effort (where 'effort' is measured as length of line)
l	is the effort for a transect or transect leg
λ	is the effort factor
n	is the number of animal groups detected
e	is the encounter rate (number of groups detected per unit length of transect, $e = n/l$
s	is the group size (number of animals observed in the group)
D	is the density (animals per unit area)
$f(0)$	is the value of the probability density function of perpendicular distances from observations to the line, evaluated at zero distance.

Subscript i is used to refer to the transect, i=1..k, and subscript j refers to the leg within the transect, j=1..m_i. Thus l_{ij} is the actual effort used for the jth leg of the ith transect (Figure 4.2). Subscript Y refers to the observation within a leg, Y=1..n_{ij}. So that s_{ijY} refers to the group size of the Yth observation of the jth leg of the ith transect.

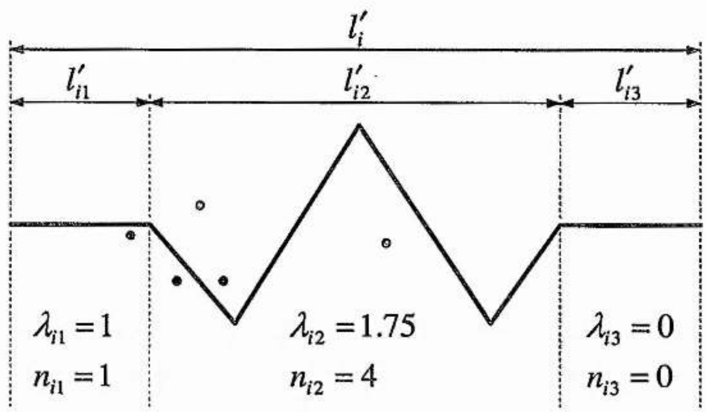


Figure 4.2: Example notation, with constant effort factor over an adaptive section. The thick zigzag line of length l_i signifies the actual effort whilst the total nominal effort would be the straight line, l'_i, for the section shown.

Nominal values refer to the values expected if a conventional straight-line transect is followed. Nominal effort is signified with a dash, such as L', the total nominal effort, whereas the corresponding actual effort is L. So, for example, the expected number of observations, if only the nominal effort had been used, is represented by $E(n \mid L')$. The same approach is also used for both the expected encounter rate and expected group size if only the nominal effort had been used, giving, for example, $E(e_i \mid l'_i)$ and $E(s_i \mid l'_i)$.

4.2.3 Assumptions

In deriving the estimating equations the following conventional line transect assumptions are made:

i) Probability of detection on the line $g(0)$, is 1.

ii) There is no size bias (the probability of detection is independent of the group size).

iii) There is no responsive movement of animals in advance of detection, and any non-responsive movement is slow relative to the speed of the observers.

These assumptions can be weakened or removed using similar strategies as for conventional line transect sampling. In addition the following assumptions are made specifically for these adaptive line transect methods:

iv) The expected encounter rate for an adaptive track is the same as the expected encounter rate for the corresponding nominal track.

v) The expected group size for an observation on an adaptive track is the same as the expected group size for an observation when following the corresponding nominal track.

vi) Conditional on the location of the actual (as distinct from the nominal) track line, each observation is an independent event. That is, the probability of an observation is only a function of its perpendicular distance from the actual line (although the position of the line itself may depend on past observations).

Approaches to dealing with heterogeneity in the detection function estimate, and thus weakening assumption vi), are explored in section 4.3.

4.2.4 Effort Factor Calculation

The effort factor, λ, is the ratio of the actual effort to the nominal effort, so that

$$\lambda = \frac{\text{Actual Effort}}{\text{Nominal Effort}}$$

Thus the effort factor for the jth leg of the ith transect is

$$\lambda_{ij} = \frac{l_{ij}}{l'_{ij}} \tag{4.1}$$

The effort factor is calculated as a function of the remaining effort available and the expected number of times the effort will be increased (i.e. the expected number of times the survey will adapt for the remaining survey).

Let

$L_E(t)$ be the total excess effort remaining at time t. This can be measured in units of time or distance.

$L'_R(t)$	be the nominal effort required to complete the survey at time t.
$L_U(t)$	be the total effort used at time t.

Thus $L_E(t)$ is calculated as total effort available at the start of the survey, less the actual effort used up to time t, less the nominal effort required to complete the survey (without any further adaptive effort). So that

$$L_E(t) = L - L'_R(t) - L_U(t)$$

Let ξ be the expected number of times the effort will increase above the nominal level for the remainder of the survey. Then the increase in effort, following an observation, is given by the excess effort available divided by the expected number of times the effort will increase plus one (for the current increase). So the increase in effort for a leg is given by

$$l_{ij} - l'_{ij} = \frac{L_E(t)}{1+\xi}$$

by definition, $l_{ij} = l'_{ij} \cdot \lambda_{ij}$, so

$$l'_{ij} \cdot (\lambda_{ij} - 1) = \frac{L_E(t)}{1+\xi}$$

and the effort factor is given by

$$\lambda_{ij} = 1 + \frac{L_E(t)}{l'_{ij} \cdot (1+\xi)} \tag{4.2}$$

If each effort increase is applied for the same fixed distance along the nominal trackline, then ξ can be calculated from an estimate of the trigger rate (the expected number of times the trigger condition will be met per unit of effort). Let l'_Z be the nominal effort over which the effort increase occurs; and γ be an estimate of the trigger rate (γ could be obtained from previous survey data or be a best guess provided by the user). If the trigger condition is a single detection, then γ is the expected encounter rate. Then

$$\xi = \gamma \cdot \{L'_R(t) - \xi \cdot l'_Z\}$$

so

$$\xi = \frac{\gamma \cdot L'_R(t)}{1 + \gamma \cdot l'_Z} \tag{4.3}$$

Thus when effort is increased over a fixed distance along the nominal trackline l'_Z (i.e. $l'_{ij} = l'_Z$ for all i, j), then the effort factor is calculated as

$$\lambda_{ij} = 1 + \frac{\left(L - L_U(t) - L'_R(t)\right)}{l'_Z \cdot \left(1 + \frac{\gamma \cdot L'_R(t)}{1 + \gamma \cdot l'_Z}\right)} \tag{4.4}$$

4.2.5 Estimating Equations

Conceptually, we estimate animal density separately for each transect line, using formulae from conventional line transect sampling. To avoid bias arising from concentrating more effort in areas of high density, weighted means of encounter rate and group size are found, weighting by the reciprocal of the effort factor. To simplify the methodology, we assume that $f(0)$ is independent of animal density, and use a single pooled estimate of $f(0)$. Later in the chapter we extend the approach to allow the modelling of heterogeneity in $f(0)$.

Density Estimate

Conventional Line Transect Estimate

The density for conventional line transect sampling, assuming detection on the line is certain, is given by (Buckland *et al.*, 2001: p54)

$$D = \frac{E(n) \cdot f(0) \cdot E(s)}{2L'} \tag{4.5}$$

Assuming $f(0)$ is constant across transects, the density corresponding to the ith transect is

$$D_i = \frac{E(n_i) \cdot f(0) \cdot E(s_i)}{2l'_i} \qquad ,i = 1..k \tag{4.6}$$

Let $\hat{f}(0)$ be a single pooled estimate of $f(0)$ for the survey. Then, replacing the parameters by their estimators, an estimate of the density for the ith transect is

$$\hat{D}_i = \frac{\hat{E}(n_i) \cdot \hat{f}(0) \cdot \hat{E}(s_i)}{2l'_i} \qquad ,i = 1..k \tag{4.7}$$

So, from Buckland *et al.* (2001: p80), an estimate of the overall density is

$$\hat{D} = \frac{1}{L'}\sum_{i=1}^{k} l_i' \hat{D}_i \tag{4.8}$$

To estimate the variance of the density for a conventional line transect survey, if the density components are estimated on a per transect basis, then from Buckland *et al.* (2001: p80)

$$\hat{V}(\hat{D}) = \frac{1}{L' \cdot (k-1)} \sum_{i=1}^{k} \left\{ l_i' \cdot \left(\hat{D}_i - \hat{D}\right)^2 \right\}$$

However, the estimate $\hat{f}(0)$ is made by pooling data across transects. Dividing out this common estimate, we have

$$\hat{V}\left(\frac{\hat{D}}{\hat{f}(0)}\right) = \frac{1}{L' \cdot (k-1)} \sum_{i=1}^{k} \left\{ l_i' \cdot \left(\frac{\hat{D}_i}{\hat{f}(0)} - \frac{\hat{D}}{\hat{f}(0)}\right)^2 \right\}$$

So an estimate of the variance of the density estimate has two components

$$\hat{V}\left(\frac{\hat{D}}{\hat{f}(0)}\right)$$

and

$$\hat{V}\left(\hat{f}(0)\right)$$

Using the delta method (Seber 1982: p5-7), an estimate of the variance of the density estimate is given by

$$\hat{V}(\hat{D}) = \hat{D}^2 \cdot \left[\frac{\hat{V}(\hat{H})}{\hat{H}^2} + \frac{\hat{V}\{\hat{f}(0)\}}{\{\hat{f}(0)\}^2} \right] \tag{4.9}$$

where

$$\hat{H} = \frac{\hat{D}}{\hat{f}(0)}$$

and

$$\hat{V}(\hat{H}) = \frac{1}{L' \cdot (k-1)} \sum_{i=1}^{k} \left\{ l_i \cdot \left(\hat{H}_i - \hat{H}\right)^2 \right\}$$

with

$$\hat{H}_i = \frac{\hat{D}_i}{\hat{f}(0)}$$

Adaptive Line Transect Density Estimate

In an adaptive line transect survey, we systematically place greater effort in areas of higher density. Thus the overall encounter rate is a biased estimate of the expected encounter rate for a conventional survey design. In the following sections we develop downweighted estimates of the sample and group size to account for this bias. For the time being, let $\hat{E}(n_i \mid l_i')$ and $\hat{E}(s_i \mid l_i')$ be, respectively, estimates of the number of observations and group size if the nominal track line had been followed. Then replacing the parameters in equation 4.6 by their adaptive estimators, we have:

$$\hat{D}_i = \frac{\hat{E}(n_i \mid l_i') \cdot \hat{f}(0) \cdot \hat{E}(s_i \mid l_i')}{2l_i'} \tag{4.10}$$

For conventional surveys, where the effort factor is 1, these estimators are simply n_i and, assuming no size bias, $\bar{s}_i$ respectively, where $\bar{s}_i$ is the mean size of groups detected on the transect. Derivation of the estimators is explained in the sections that follow.

This estimate of the adaptive line transect density is used in the same manner as the conventional transect estimator to get estimates $\hat{D}$ and $\hat{V}(\hat{D})$, using equations 4.8 and 4.9.

Effort

By definition the nominal effort for the j^{th} leg of the i^{th} transect is

$$l_{ij}' = l_{ij} / \lambda_{ij}$$

with the nominal transect effort and nominal total survey effort given by

$$l_i' = \sum_{j=1}^{m_i} l_{ij}'$$

and

$$L' = \sum_{i=1}^{k} l_i'$$

respectively.

Number of Observations

An estimate of the number of observations if only the nominal effort had been used for the j^{th} leg of the i^{th} transect is given by

$$\hat{E}(n_{ij} \mid l'_{ij}) = \frac{n_{ij}}{\lambda_{ij}}$$

with the corresponding transect and survey estimates of

$$\hat{E}(n_i \mid l'_i) = \sum_{j=1}^{m_i} \hat{E}(n_{ij} \mid l'_{ij}) \tag{4.11}$$

and

$$\hat{E}(n \mid L') = \sum_{i=1}^{k} \hat{E}(n_i \mid l'_i) \tag{4.12}$$

An estimate of the variance of estimated expected number of observations if only the nominal effort had been used is given by

$$\hat{V}\left\{\hat{E}(n \mid L')\right\} = \frac{L'}{k-1} \cdot \sum_{i=1}^{k} \left[l'_i \cdot \left\{\hat{E}(n_i \mid l'_i) - \hat{E}(n \mid L')\right\}^2 \right] \tag{4.13}$$

Encounter Rate

The encounter rate for the j^{th} leg of the i^{th} transect is given by

$$e_{ij} = n_{ij} / l_{ij}$$

So from assumption iv) an estimate of the expected encounter rate if only the nominal effort had been used for the j^{th} leg of the i^{th} transect, is given by

$$\hat{E}(e_{ij} \mid l') = \frac{n_{ij}}{l_{ij}} = \frac{\hat{E}(n_{ij} \mid l'_{ij})}{l'_{ij}}$$

Using weighted averages, an estimate of the expected encounter rate if only the nominal effort was used for the i^{th} transect, is

$$\hat{E}(e_i \mid l'_i) = \frac{\sum_{j=1}^{m_i} \left\{ l'_{ij} \cdot \hat{E}(e_{ij} \mid l'_{ij}) \right\}}{\sum_{j=1}^{m_i} l'_{ij}} = \frac{\sum_{j=1}^{m_i} \left\{ \hat{E}(n_{ij} \mid l'_{ij}) \right\}}{\sum_{j=1}^{m_i} l'_{ij}} = \frac{\hat{E}(n_i \mid l'_i)}{l'_i} \tag{4.14}$$

and an estimate of the survey encounter rate if only the nominal effort was used, is

$$\hat{E}(e \mid L') = \frac{\sum_{i=1}^{k} \left\{ l'_i \cdot \hat{E}(e_i \mid l'_i) \right\}}{\sum_{i=1}^{k} l'_i} = \frac{\sum_{i=1}^{k} \left\{ \hat{E}(n_i \mid l'_i) \right\}}{\sum_{i=1}^{k} l'_i} = \frac{\hat{E}(n \mid L')}{L'} \tag{4.15}$$

Thus an estimate of the variance of the expected survey encounter rate if only the nominal effort had been used is

$$\hat{V}\{\hat{E}(e \mid L')\} = \hat{V}\left\{\frac{\hat{E}(n \mid L')}{L'}\right\} = \frac{\hat{V}\{\hat{E}(n \mid L')\}}{(L')^2} \tag{4.16}$$

Group size

The mean observed group size for the jth leg of the ith transect is

$$\bar{s}_{ij} = \frac{\sum_{Y=1}^{n_{ij}} s_{ijY}}{n_{ij}}$$

Assuming there is no size biased detection, and that the expected group size for a leg following an adaptive track is the same as the expected group size when following the corresponding nominal track, assumption (v), i.e.

$$E(s_{ij}) = E(s_{ij} \mid l'_{ij})$$

then an estimate of the expected school size for the jth leg of the ith transect is

$$\hat{E}(s_{ij}) = \bar{s}_{ij}$$

so an estimate of the expected group size for the jth leg of the ith transect using nominal effort is

$$\hat{E}(s_{ij} \mid l'_{ij}) = \bar{s}_{ij} = \frac{\sum_{Y=1}^{n_{ij}} s_{ijY}}{n_{ij}}$$

and the expected total number of animals observed for the jth leg of the ith transect following a nominal trackline is

$$\hat{E}(n_{ij} \mid l'_{ij}) \cdot \hat{E}(s_{ij} \mid l'_{ij}) = \frac{n_{ij}}{\lambda_{ij}} \cdot \frac{\sum_{Y=1}^{n_{ij}} s_{ijY}}{n_{ij}} = \frac{\sum_{Y=1}^{n_{ij}} s_{ijY}}{\lambda_{ij}}$$

Using weighted averages, an estimate of the expected group size for the ith transect, if only the nominal effort had been used, is given by

$$\hat{E}(s_i \mid l'_i) = \frac{\sum_{j=1}^{m_i} \{\hat{E}(n_{ij} \mid l'_{ij}) \cdot \hat{E}(s_{ij} \mid l'_{ij})\}}{\sum_{j=1}^{m_i} \hat{E}(n_{ij} \mid l'_{ij})} = \frac{\sum_{j=1}^{m_i} \{\hat{E}(n_{ij} \mid l'_{ij}) \cdot \hat{E}(s_{ij} \mid l'_{ij})\}}{\hat{E}(n_i \mid l'_i)} \tag{4.17}$$

Similarly an estimate of the expected group size, if only the nominal effort had been used, is

$$\hat{E}(s \mid L') = \frac{\sum_{i=1}^{k}\left\{\hat{E}(n_i \mid l_i') \cdot \hat{E}(s_i \mid l_i')\right\}}{\sum_{i=1}^{k}\hat{E}(n_i \mid l_i')} = \frac{\sum_{i=1}^{k}\left\{\hat{E}(n_i \mid l_i') \cdot \hat{E}(s_i \mid l_i')\right\}}{\hat{E}(n \mid L')} \tag{4.18}$$

An estimate of the variance of the expected group size, if only the nominal effort had been used, is

$$\hat{V}\left\{\hat{E}(s \mid L')\right\} = \frac{1}{\hat{E}(n \mid L') \cdot (k-1)} \sum_{i=1}^{k}\left[\hat{E}(n_i \mid l_i') \cdot \left\{\hat{E}(s_i \mid l_i') - \hat{E}(s \mid L')\right\}^2\right] \tag{4.19}$$

f(0)

It is assumed there is no correlation between density and $f(0)$ and so observation data are pooled across all transects to produce a single estimate of $f(0)$ using conventional techniques (see, for example, Buckland *et al.*, 2001). This approach does not allow for heterogeneity between groups in the probability of detection due to group size, weather conditions, etc. In the next section, Modelling Heterogeneity in f(0), we explore approaches to deal with this.

4.3 Modelling Heterogeneity in *f*(0)

4.3.1 Introduction

The above methods do not allow for heterogeneity between groups in the probability of detection due to group size, weather conditions, etc. In practice, adaptive effort is more likely to be triggered in good sighting conditions, and so the probability of detection on the adaptive leg may be enhanced. Since such observations will be over-represented in the sample, the $f(0)$ estimate (which is the reciprocal of the estimated effective strip half-width) will be negatively biased.

We start the section by discussing basic tests to detect heterogeneity in $f(0)$; then consider an ad hoc method to account for heterogeneity by weighting the observation data used to estimate $f(0)$; and finish by developing a new estimator that allows the inclusion of covariates in the $f(0)$. This new estimator builds on recent developments in distance sampling to address unequal probability of detection (see for example Borchers *et al.*, 1998a; Borchers *et al.*, 1998b; Marques, 2001).

4.3.2 Detecting Heterogeneity in *f*(0)

Sightings made when adapting are not down-weighted when estimating *f*(0), which may lead to bias in the presence of heterogeneity. For example good sighting conditions may lead to an increase in the number of adaptive triggers, which in turn may lead to increased observations on the zigzag track and so negatively bias the *f*(0) estimate. Effective strip width is wider in good sighting conditions, so that *f*(0) is smaller. It is recommended that survey results are carefully examined to check for such bias.

A possible approach is to pool the data for the observations while following a nominal track and to separately pool the data for the observations from the adaptive track. The resultant two estimates of *f*(0), $\hat{f}_N(0)$ for the nominal observations and $\hat{f}_A(0)$ for the adaptive observations, can then be tested for differences. Three potential tests are:

i) A basic z test of whether the expectations of $\hat{f}_N(0)$ and $\hat{f}_A(0)$ are the same;
ii) A χ^2 test of whether the perpendicular sighting distance distribution for observations made on the nominal track is the same as that for observations made when adapting;
iii) Akaike's Information Criteria (AIC).

The AIC approach is applied by comparing the sum of the AICs for modelling $\hat{f}_N(0)$ and $\hat{f}_A(0)$ separately with the AIC for modelling *f*(0) using sightings pooled across the two survey modes. If the AIC value for the pooled model is less than the sum of the other two AICs then this suggests that a single model fits the data better than two separate models. So, as a rough guide, if the AIC value for the pooled model is greater than the sum of the other two AICs, then this could be taken as a sign of heterogeneity.

Initial simulation trials suggested these tests have low power, and further investigation is required.

Weighting the Maximum Likelihood Estimate of *f*(0)

We can seek to model the heterogeneity. However, if we have not measured the relevant covariates, or if probability of detection of further animals changes following an observation (because observers become more alert or they continue to watch detected animals), this approach may not be wholly effective.

Here we use the simplistic approach of downweighting the influence of adaptive observations on the maximum likelihood estimate of $f(y)$. Adopting the principle that a single observation at distance y from the line with effort factor 1 should have the same contribution to the likelihood as λ observations at distance y with effort factor λ, we obtain a modified likelihood.

Let the density function be expressed by (Buckland *et al.* 2001: p59)

$$f(y) \doteq \frac{\alpha(y)}{\beta} \cdot \left[1 + \sum_{j=1}^{m} a_j \cdot p_j(y_s)\right]$$

where

y	is the perpendicular distance from the trackline
$\alpha(y)$	is a parametric key, containing k parameters
$p_j(y_s)$	is either a simple polynomial; the j$^{\text{th}}$ Hermite polynomial; or a cosine series
y_s	is a scale of y
a_j	is the adjustment term
β	is a normalising function of the parameters

So to weight the observations such that λ observations at distance y with effort factor λ make the same contribution as 1 observation at distance y with effort factor 1, we modify the $f(y)$ estimate by raising to the power $1/\lambda_{ij}$ each $f(y)$ component in the maximum likelihood estimate. Thus the modified likelihood is given by

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{i=1}^{n} \{f(y_i)\}^{\frac{1}{\lambda_i}} \tag{4.20}$$

where

y_i	is the perpendicular distance from the line of the i^{th} observation, $i = 1,...,n$
λ_i	is the effort factor corresponding to the i^{th} observation
θ	represents the parameters of $f(y_i)$. $\theta_1,...,\theta_k$ are the parameters of the key function and $\theta_{k+j} = a_j$, $j=1,...,m$ are the parameters of the adjustment terms

and the log-likelihood is given by

$$\begin{aligned}\log_e[\mathcal{L}(\underline{\theta})] &= \sum_{i=1}^{n}\frac{1}{\lambda_i}\log_e[f(y_i)] \\ &= \sum_{i=1}^{n}\frac{1}{\lambda_i}\log_e[f(y_i)\cdot\beta] - n\cdot\log_e\beta\cdot\sum_{i=1}^{n}\frac{1}{\lambda_i}\end{aligned} \tag{4.21}$$

This modified likelihood function is then maximised with respect to the parameters of $f(.)$. For example, consider the half-normal model with ungrouped data and no truncation, so θ is the scalar σ^2 and $f(y) = e^{-y^2/2\sigma^2}$, $y \geq 0$. Then a weighted estimate of $f(0)$ is given by

$$\hat{f}_w(0) = \sqrt{2/\pi\hat{\sigma}_w^{\,2}} \tag{4.22}$$

where

$$\hat{\sigma}_w^{\,2} = \frac{\sum_{i=1}^{n} y_i^2/\lambda_i}{\sum_{i=1}^{n} 1/\lambda_i} = \frac{\sum_{i=1}^{n} y_i^2/\lambda_i}{\hat{E}(n\,|\,L')}$$

and the variance estimate is given by

$$\hat{V}\{\hat{f}_w(0)\} = \frac{\{\hat{f}_w(0)\}^2}{2\cdot\hat{E}(n\,|\,L')} \tag{4.23}$$

Derivation of Weighted MLE for Half-Normal Detection Function

This example considers the half-normal detection function, for line transects with ungrouped data and no truncation. Please note, that for this derivation, subscript i is used to refer to the i^{th} observation of the n observations in the survey.

The half-normal detection function is given by

$$g(x) = \exp\left(-x^2/2\sigma^2\right) \quad , \quad 0 \le x \le \infty$$

With no truncation, the density function of detection distances is

$$f(x) = g(x)/\mu$$

where

$$\mu = \int_0^\infty g(x)\,dx = \int_0^\infty \exp\left(-x^2/2\sigma^2\right)dx = \sqrt{\frac{\pi\sigma^2}{2}}$$

Using the weighted likelihood approach then for n detections the likelihood function is given by

$$\mathcal{L} = \prod_{i=1}^{n}\{f_w(0)\}^{\frac{1}{\lambda_i}} = \prod_{i=1}^{n}\{g_w(x_i)/\mu_w\}^{\frac{1}{\lambda_i}} = \prod_{i=1}^{n}\left\{\frac{\exp\left(-x_i^2/2\sigma_w^2\right)}{\sqrt{\pi\sigma_w^2/2}}\right\}^{\frac{1}{\lambda_i}}$$

Thus the log-likelihood is given by

$$logl = \sum_{i=1}^{n}\frac{1}{\lambda_i}\log_e\left(\frac{\exp\left(-x_i^2/2\sigma_w^2\right)}{\sqrt{\pi\sigma_w^2/2}}\right) = \sum_{i=1}^{n}\frac{-x_i^2}{2\lambda_i\sigma_w^2} - \sum_{i=1}^{n}\frac{\log_e\left(\sqrt{\pi\sigma_w^2/2}\right)}{\lambda_i}$$

Differentiating with respect to σ_w^2 gives

$$\frac{d\,logl}{d\sigma_w^2} = \sum_{i=1}^{n}\left\{\left(\frac{-x_i^2}{2\lambda_i}\right)\cdot(-1)\cdot\left(\frac{1}{\sigma_w^4}\right)\right\} -$$

$$\sum_{i=1}^{n}\left\{\left(\frac{1}{\lambda_i}\right)\cdot\left(\frac{1}{\sqrt{\pi\sigma_w^2/2}}\right)\cdot\left(\frac{1}{2}\right)\cdot\left(\frac{1}{\sqrt{\pi\sigma_w^2/2}}\right)\cdot\left(\frac{\pi}{2}\right)\right\}$$

$$= \sum_{i=1}^{n}\frac{x_i^2}{2\lambda_i\sigma_w^4} - \sum_{i=1}^{n}\frac{1}{2\lambda_i\sigma_w^2}$$

Setting the result equal to zero gives

$$\sum_{i=1}^{n}\frac{x_i^2}{2\lambda_i\sigma_w^4} - \sum_{i=1}^{n}\frac{1}{2\lambda_i\sigma_w^2} = 0$$

so that

$$\sigma_w^2 = \frac{\sum_{i=1}^{n} x_i^2/\lambda_i}{\sum_{i=1}^{n} 1/\lambda_i} \tag{4.24}$$

and so σ_w^2 can be estimated by

$$\hat{\sigma}_w^2 = \frac{\sum_{i=1}^{n} x_i^2/\lambda_i}{\sum_{i=1}^{n} 1/\lambda_i} = \frac{\sum_{i=1}^{n} x_i^2/\lambda_i}{\hat{E}(n \mid L')} \tag{4.25}$$

and, if the probability of detection on the line is certain, then

$$\hat{f}_w(0) = g_w(0)/\hat{\mu}_w = 1/\hat{\mu}_w = \sqrt{2/\pi\hat{\sigma}_w^2} \tag{4.26}$$

Now

$$\begin{aligned}
\frac{\partial\, logl^2}{\partial^2 \sigma_w^2} &= \left\{\sum_{i=1}^{n}\left(\frac{x_i^2}{2\lambda_i}\right)\cdot(-2)\cdot\left(\frac{1}{\sigma_w^6}\right)\right\} - \left\{\sum_{i=1}^{n}\left(\frac{1}{2\lambda_i}\right)\cdot(-1)\cdot\left(\frac{1}{\sigma_w^4}\right)\right\} \\
&= \frac{-1}{\sigma_w^6}\sum_{i=1}^{n}\frac{x_i^2}{\lambda_i} + \frac{1}{2\sigma_w^4}\sum_{i=1}^{n}\frac{1}{\lambda_i}
\end{aligned}$$

Thus the Fisher Information function $I(\theta)$ is given by

$$\begin{aligned}
I(\sigma_w^2) &= E\left(-\partial\, logl^2/\partial^2\, \sigma_w^2\right) \\
&= E\left(-\left\{\frac{-1}{\sigma_w^6}\sum_{i=1}^{n}\frac{x_i^2}{\lambda_i} + \frac{1}{2\sigma_w^4}\sum_{i=1}^{n}\frac{1}{\lambda_i}\right\}\right) \\
&= \frac{\left(\sum_{i=1}^{n}\frac{x_i^2}{\lambda_i}\right)}{\sigma_w^6} - \frac{1}{2\sigma_w^4}\sum_{i=1}^{n}\frac{1}{\lambda_i}
\end{aligned}$$

So using equation 4.24 this simplifies to

$$\begin{aligned}
I(\sigma_w^2) &= \frac{\left(\sum_{i=1}^{n}\frac{1}{\lambda_i}\right)\cdot\sigma_w^2}{\sigma_w^6} - \frac{1}{2\sigma_w^4}\sum_{i=1}^{n}\frac{1}{\lambda_i} \\
&= \frac{1}{\sigma_w^4}\sum_{i=1}^{n}\frac{1}{\lambda_i} - \frac{1}{2\sigma_w^4}\sum_{i=1}^{n}\frac{1}{\lambda_i} \\
&= \frac{1}{2\sigma_w^4}\sum_{i=1}^{n}\frac{1}{\lambda_i}
\end{aligned}$$

So $V\left(\hat{\sigma}_w^2\right)$ can be approximated by

$$V\left(\hat{\sigma}_w^2\right) \approx \frac{1}{I\left(\sigma_w^2\right)} = \frac{1}{\dfrac{1}{2\sigma_w^4}\displaystyle\sum_{i=1}^{n}\frac{1}{\lambda_i}} = \frac{2\sigma_w^4}{\displaystyle\sum_{i=1}^{n}\frac{1}{\lambda_i}}$$

and an estimate of $V\left(\hat{\sigma}_w^2\right)$ is given by

$$\hat{V}\left(\hat{\sigma}_w^2\right) = \frac{2\hat{\sigma}_w^4}{\displaystyle\sum_{i=1}^{n}\frac{1}{\lambda_i}} = \frac{2\hat{\sigma}_w^4}{\hat{E}(n \mid L')} \tag{4.27}$$

Using the delta method (Seber, 1982: p5-7)

$$V\left\{\hat{f}_w(0)\right\} \approx V\left(\hat{\sigma}_w^2\right)\cdot\left(\frac{\partial f_w(0)}{\partial \sigma_w^2}\right)^2$$

So an estimate of the variance of $\hat{f}_w(0)$ is given by

$$\begin{aligned}
\hat{V}\left\{\hat{f}_w(0)\right\} &\approx \hat{V}\left(\hat{\sigma}_w^2\right)\cdot\left(\frac{\partial}{\partial \hat{\sigma}_w^2}\sqrt{2/\pi\hat{\sigma}_w^2}\right)^2 \\
&= \left(\frac{2\hat{\sigma}_w^4}{\hat{E}(n \mid L')}\right)\cdot\left[\left\{\left(\frac{1}{2}\right)\left(\frac{2}{\pi\hat{\sigma}_w^2}\right)^{\frac{-1}{2}}\cdot\left(\frac{2}{\pi}\right)\cdot(-1)\cdot\left(\frac{1}{\hat{\sigma}_w^4}\right)\right\}^2\right] \\
&= \left(\frac{2\hat{\sigma}_w^4}{\hat{E}(n \mid L')}\right)\cdot\left[\left\{\left(\frac{-1}{\pi\hat{\sigma}_w^4}\right)\cdot\left(\sqrt{\frac{\pi\hat{\sigma}_w^2}{2}}\right)\right\}^2\right] \\
&= \left(\frac{2\hat{\sigma}_w^4}{\hat{E}(n \mid L')}\right)\cdot\left(\frac{-1}{\pi\hat{\sigma}_w^4}\right)^2\cdot\left(\frac{\pi\hat{\sigma}_w^4}{2}\right) \\
&= \left(\frac{2\hat{\sigma}_w^4}{\hat{E}(n \mid L')}\right)\cdot\left(\frac{1}{2\pi\hat{\sigma}_w^6}\right) \\
&= \frac{1}{\pi\hat{\sigma}_w^2\cdot\hat{E}(n \mid L')}
\end{aligned}$$

Substituting $\hat{f}_w(0) = \sqrt{2/\pi\hat{\sigma}_w^2}$ from equation 4.25 gives

$$\hat{V}\left\{\hat{f}_w(0)\right\} = \frac{\left\{\hat{f}_w(0)\right\}^2}{2\cdot\hat{E}(n \mid L')} \tag{4.28}$$

4.3.3 Including Covariates in *f*(0)

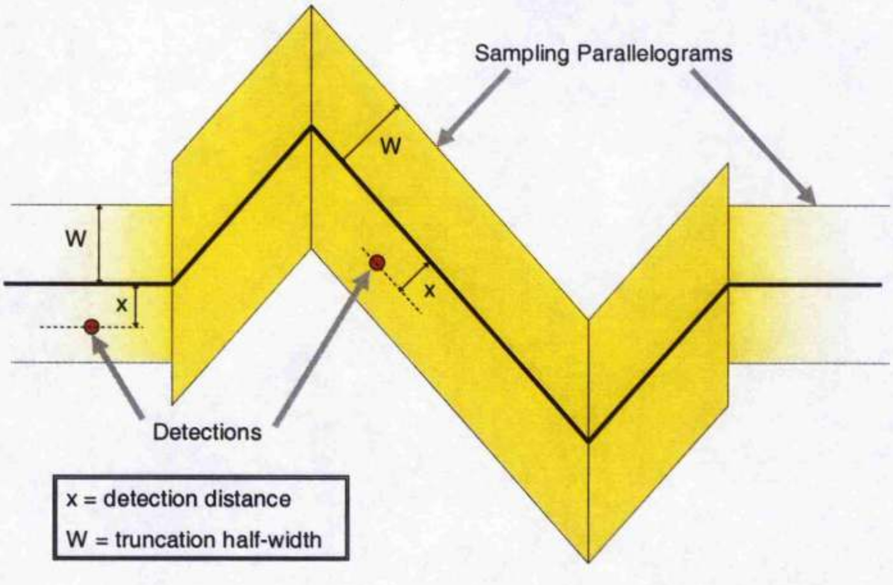


Figure 4.3: Sampling strip parallelograms, shown for a zigzag adaptive pattern. Animals are sampled within the parallelogram out as far as the truncation half-width, *W*.

Another way to deal with the heterogeneity is to use covariates in the estimation of the detection function. Here we expand the PB method to accommodate a probability of inclusion.

Defining

N	is the number of (individual) animals in the survey region
N_g	is the number of animal groups in the survey region
A	is the area of the survey region
δ	is the number of animals in the surveyed area
τ	is the number of animal groups in the surveyed area

Consider the sampling strip out to truncation half-width, *W*, on either side of the trackline. Then for the PB method, each transect leg can be considered to have a *sampling strip parallelogram*. For example, in Figure 4.3 sampling strip parallelograms are shown for a zigzag adaptive pattern. Thus using a Horvitz-Thompson-based estimator (Horvitz and Thompson, 1952), an estimate of the number of animals in the j^{th} sampling strip parallelogram of the i^{th} transect is given by

$$\hat{\delta}_{ij} = \sum_{Y=1}^{n_{ij}} \frac{s_{ijY}}{\hat{p}_{ijY}}$$

where

n_{ij}	is the number of animal groups (observations) on the j^{th} transect leg of the i^{th} transect.
s_{ijY}	is the number of animals in the Y^{th} group (observation) of the j^{th} transect leg of the i^{th} transect
$\hat{p}_{ijY}$	is an estimate of the probability of detection of the Y^{th} group on the j^{th} transect leg of the i^{th} transect, within the strip half-width.

So an estimate of the expected number of animals in the j^{th} sampling parallelogram of the i^{th} transect, if the survey had followed the nominal trackline is

$$\hat{E}\left(\hat{\delta}_{ij} \mid l'_{ij}\right) = \frac{\hat{\delta}_{ij}}{\lambda_{ij}} = \frac{1}{\lambda_{ij}} \sum_{Y=1}^{n_{ij}} \frac{s_{ijY}}{\hat{p}_{ijY}} \tag{4.29}$$

where

λ_{ij}	is the effort factor for the j^{th} transect leg of the i^{th} transect
l'_{ij}	is the nominal effort for the j^{th} transect leg of the i^{th} transect

and an estimate of the expected number of animals in the surveyed area for the i^{th} transect, had the survey only followed the nominal trackline, is

$$\hat{E}\left(\hat{\delta}_{i} \mid l'_{i}\right) = \sum_{j=1}^{m_i} \hat{E}\left(\hat{\delta}_{ij} \mid l'_{ij}\right) = \sum_{j=1}^{m_i} \frac{1}{\lambda_{ij}} \sum_{Y=1}^{n_{ij}} \frac{s_{ijY}}{\hat{p}_{ijY}} \tag{4.30}$$

where

m_i	is the number of transect legs in the i^{th} transect
l'_i	is the nominal effort for the i^{th} transect

and finally, an estimate of the expected number of animals in the surveyed area, had the survey only followed the nominal trackline, is

$$\hat{E}\left(\hat{\delta} \mid L'\right) = \sum_{i=1}^{k} \hat{E}\left(\hat{\delta}_{i} \mid l'_{i}\right) = \sum_{i=1}^{k} \sum_{j=1}^{m_i} \frac{1}{\lambda_{ij}} \sum_{Y=1}^{n_{ij}} \frac{s_{ijY}}{\hat{p}_{ijY}} \tag{4.31}$$

where

k	is the number of transects in the survey
L'	is the total nominal effort for the survey

The surveyed area, assuming the survey had only followed the nominal trackline, A'_s, is given by

$$A'_s = 2WL' \tag{4.32}$$

Thus the estimate of the animal density is given by

$$\begin{aligned}\hat{D}_{HT} &= \frac{A}{A_s}\hat{E}\left(\hat{\delta} \mid L'\right) \\ &= \frac{A}{2WL'}\hat{E}\left(\hat{\delta} \mid L'\right) \\ &= \frac{A}{2WL'}\sum_{i=1}^{k}\sum_{j=1}^{m_i}\frac{1}{\lambda_{ij}}\sum_{Y=1}^{n_{ij}}\frac{s_{ijY}}{\hat{p}_{ijY}}\end{aligned} \tag{4.33}$$

From 3.27, an estimate of p_{ijY} is given by

$$\hat{p}_{ijY} = \frac{1}{W}\int_0^W \hat{g}\left(x \mid \mathbf{z}_{ijY}\right)dx$$

where

$\hat{g}\left(x \mid \mathbf{z}_{ijY}\right)$ is an estimate of the probability of detection, at perpendicular distance x from the trackline, for the X$^{\text{th}}$ object of the j$^{\text{th}}$ transect leg of the i$^{\text{th}}$ transect, with associated covariates $\mathbf{z}_{ijY}$.

Substituting this in equation 4.33 gives

$$\begin{aligned}\hat{D}_{HT} &= \frac{A}{2WL'}\sum_{i=1}^{k}\sum_{j=1}^{m_i}\frac{1}{\lambda_{ij}}\sum_{Y=1}^{n_{ij}}\frac{W \cdot s_{ijY}}{\int_0^W \hat{g}\left(x \mid \mathbf{z}_{ijY}\right)dx} \\ &= \frac{A}{2L'}\sum_{i=1}^{k}\sum_{j=1}^{m_i}\frac{1}{\lambda_{ij}}\sum_{Y=1}^{n_{ij}}\frac{s_{ijY}}{\int_0^W \hat{g}\left(x \mid \mathbf{z}_{ijY}\right)dx}\end{aligned} \tag{4.34}$$

From 3.29, if detection on the line is certain, so that $g\left(0 \mid \mathbf{z}_{ijY}\right) = 1$, then this further simplifies to

$$\hat{D}_{HT} = \frac{A}{2L'}\sum_{i=1}^{k}\sum_{j=1}^{m_i}\frac{1}{\lambda_{ij}}\sum_{Y=1}^{n_{ij}} s_{ijY} \cdot \hat{f}\left(0 \mid \mathbf{z}_{ijY}\right) \tag{4.35}$$

We consider the variance of the density estimate to have two main components; the variance from the encounter rate and the variance from the combined group sized and detection function estimate for observations, the summed $s_{ijY} / \hat{p}_{ijY}$ component.

So, using the delta method (Seber, 1982: p5-7), an estimate of the variance is given by

$$\hat{V}\left(\hat{D}_{HT}\right) = \hat{D}_{HT}^2 \cdot \left[\left\{\hat{cv}\left(\hat{E}[e \mid L']\right)\right\}^2 + \left\{\hat{cv}\left(\sum_{i=1}^{k}\sum_{j=1}^{m_i}\sum_{Y=1}^{n_{ij}} \frac{s_{ijY}}{\hat{p}_{ijY}}\right)\right\}^2 \right] \tag{4.36}$$

For the encounter rate the estimate of the expected nominal value, $\hat{E}(e \mid L')$ is estimated using equation 4.15, with variance estimated by equation 4.16. Assuming that the observations are independent of one another with respect to group size and probability of detection, and group size is a covariate in the p_{ijY} estimate, then the coefficient of variation (cv) of $\sum_{i=1}^{k}\sum_{j=1}^{m_i}\sum_{Y=1}^{n_{ij}} s_{ijY} / \hat{p}_{ijY}$ can be estimated by bootstrapping (Efron and Tibshirani, 1993).

The density of groups, $D_{g.HT}$, can be estimated by replacing the s_{ijY} by 1 in all equations, so that equation 4.34 becomes

$$\hat{D}_{g.HT} = \frac{A}{2L'}\sum_{i=1}^{k}\sum_{j=1}^{m_i}\frac{1}{\lambda_{ij}}\sum_{Y=1}^{n_{ij}}\frac{1}{\int_0^W \hat{g}\left(x \mid \mathbf{z}_{ijY}\right)dx} \tag{4.37}$$

and if detection on the line is certain, then

$$\hat{D}_{g.HT} = \frac{A}{2L'}\sum_{i=1}^{k}\sum_{j=1}^{m_i}\frac{1}{\lambda_{ij}}\sum_{Y=1}^{n_{ij}}\hat{f}\left(0 \mid \mathbf{z}_{ijY}\right) \tag{4.38}$$

Following the estimate of animal density, an estimate of the group density variance is given by

$$\hat{V}\left(\hat{D}_{HT}\right) = \hat{D}_{HT}^2 \cdot \left[\left\{\hat{cv}\left(\hat{E}[e \mid L']\right)\right\}^2 + \left\{\hat{cv}\left(\sum_{i=1}^{k}\sum_{j=1}^{m_i}\sum_{Y=1}^{n_{ij}} \frac{1}{\hat{p}_{ijY}}\right)\right\}^2 \right] \tag{4.39}$$

where the cv of $\sum_{i=1}^{k}\sum_{j=1}^{m_i}\sum_{Y=1}^{n_{ij}} 1/\hat{p}_{ijY}$ is again estimated by bootstrapping.

4.4 Simulation

To investigate the efficiency of the adaptive approach, a computer program was developed to simulate clustered populations and then simulate conventional and adaptive line transect surveys on these populations. The process was repeated multiple times and mean results for the two survey approaches compared. The simulation program, RATS, is described in more detail in Appendix A. For these simulations, it was assumed that animals occurred singly; i.e. the group size was always one.

Estimation of the detection function was performed using the distance sampling analysis software DISTANCE 2.2 (Laake *et al.*, 1996), using the following models: half-normal key and cosine adjustments; half-normal and Hermite polynomial adjustments; hazard-rate and cosine adjustments; hazard-rate and simple polynomial adjustments; and uniform and cosine adjustments.

Simulation was only carried out using the basic PB method for line transect surveys and there was no simulation using covariates to estimate $f(0)$.

4.4.1 Population Models

Three base types of population were simulated using the computer program: a population exhibiting complete spatial randomness (CSR); a clustered population; and a highly clustered population. Each population had an expected size of 600 and was created within a square area (population frame) of 100 by 100 units. The clustered and highly clustered populations were simulated using a Poisson cluster process (Diggle, 1983). The number of parent clusters was simulated using a Poisson(40) distribution for the clustered population and a Poisson(15) distribution for the highly clustered population. The number of animals within each parent cluster was then simulated using a Poisson(15) distribution for the clustered and a Poisson(40) distribution for the highly clustered population. For each parent cluster, the location of the centre of the cluster was simulated using a Uniform(0, 100) distribution for the vertical co-ordinate and another Uniform(0, 100) for the horizontal co-ordinate. Finally the position of each animal within each parent cluster was calculated relative to the parent cluster centre. The radial distance to each animal was simulated from a Normal(0, 4) distribution and the radial angle using a Uniform(0, 2π). If following this, the animal lay outside the population frame, the distance to the animal was wrapped around to the opposite edge, horizontally or vertically as necessary, until the animal was within the population frame. The population parameters are summarised in Table 4.1, example populations can be seen in Appendix B.

Table 4.1: Population simulation parameters for the 3 population types. For each type, these parameters give an expected population size of 600.

	Population Type		
Component	**CSR**	**Clustered**	**Highly Clustered**
Number of parent clusters	Constant(600)	Poisson(40)	Poisson(15)
X position of parent cluster centres	Uniform[0, 100]	Uniform[0, 100]	Uniform[0, 100]
Y position of parent cluster centres	Uniform[0, 100]	Uniform[0, 100]	Uniform[0, 100]
Number of objects in each parent cluster	Constant(1)	Poisson(15)	Poisson(40)
Object angle	Uniform[0, 2π]	Uniform[0, 2π]	Uniform[0, 2π]
Object radial distance	Constant(0)	Normal(0, 4)	Normal(0,4)
School size	Constant(1)	Constant(1)	Constant(1)

4.4.2 Survey Simulation

The sampling transects were run horizontally across the population frame, from left to right, starting on the left edge, so that each had a nominal length of 100 units. For each population generated, the vertical start positions of each transect were restricted to be within the range 5 to 95 units. Providing a 5 unit buffer zone at the top and bottom of the population frame, in which transects could not be located.. This minimized the potential for edge effects, where the zigzag extends out of the top or the bottom of the population frame, and so does not detect any animals. The transects were systematically spaced with a random vertical start position for the first transect. For each survey the total effort was set at 1500 units and for the adaptive surveys the nominal effort was set at 1300 units. This meant that for each conventional survey 15 sampling transects were run, across the population frame, and for each adaptive survey 13 transects were run. For each pair of conventional and adaptive surveys, the same transect start positions were used, though in the case of the adaptive surveys, only the first 13 of the 15 random start positions were used.

Thus adaptive surveys did not cover the lower section of the population frame. This is not an issue, as for each survey pair a new population was generated; the transect vertical start coordinate was randomised; there was no density gradient in the population; and so the location of transects was effectively random.

Each simulated population was sampled first using a conventional line transect survey and then using an adaptive line transect survey. The detection function was simulated using a half-normal detection function with parameter $\sigma = 0.3$, with perpendicular distances truncated at 2 units. This is effectively no truncation, as with $\sigma = 0.3$, the probability of detecting an object beyond 2 units is approximately 2.63×10^{-11}.

Each transect was traversed in horizontal steps of length 0.66 units, so that a single adaptive zigzag cycle covered approximately 2 units. For the conventional surveys a rectangular area, centred on the transect, was sampled. For the adaptive surveys, when zigzagging, the rectangle becomes a parallelogram. The truncation half-width is thus the perpendicular offset to the edge of the rectangle or parallelogram on each side of a transect.

The trigger to start adapting was a single observation in the previous 0.66 unit length of transect, after which zigzagging occurred for 12 steps (3 complete zigzag cycles), spanning 2 units of nominal line, with the angle of the zigzags adjusted appropriately for the effort factor. If there was an observation during the last step of an adaptive section, the adapting continued for another 12 steps. Each transect started in straight-line mode, irrespective of whether the survey was still adapting at the end of the previous transect. Survey parameters are summarised in Table 4.2 and example simulations for each of the 3 population types are shown in Appendix B.

Additional simulations were performed using the highly clustered populations to investigate the effects of heterogeneity in the detection function. First to simulate an increase in observer awareness, following an observation, the detection function for the adaptive surveys was changed. For the conventional surveys and the straight-line sections of the adaptive surveys a half-normal detection function with $\sigma = 0.3$ was used as before. However to simulate an increase in observer awareness, σ was increased to 0.4, on the adaptive (zigzag) sections of the adaptive survey. Second, to simulate heterogeneity introduced by changes in weather, simulation of the highly clustered population was again re-run, this time using a half-normal detection function with $\sigma = 0.15$ for the first 400 units of survey effort, reverting to $\sigma = 0.3$ for the remainder of the survey. This type of heterogeneity was applied to both the adaptive and the conventional surveys as they are equally likely to encounter bad weather. For these simulations the detection function was estimated, as a known half-normal, using equations 4.22 and 4.23.

Table 4.2: Details of parameters used to simulate the surveys.

Parameter	Value
Total effort	1500
Nominal effort	1300
Equal nominal length transects	TRUE
Detection function	Half-normal ($\sigma = 0.3$)
Truncation half-width (W)	2.0
Adaptive pattern	Zigzag
Adaptive cycles (pattern repeat)	3
Step length	0.66
Expected encounter rate	0.030

4.4.3 Simulation Results

Detailed tables summarising results for the 5 simulation runs are given in Appendix B. These tables also include estimation using the inbuilt detection function estimate within RATS, which assumes the detection function is half-normal. Thus estimates based on this do not include any model selection uncertainty. In preference results have been produced using DISTANCE 2.2 to estimate $f(0)$ from a choice of three models, representing a more realistic analysis of survey data.

For each population type, 1000 populations were generated, with both a conventional and an adaptive survey run on each population. Two additional runs of 1000 highly clustered populations were generated to test heterogeneity in the detection function, due to increased observer awareness and changes in weather conditions. The comparative efficiencies of estimators were calculated by dividing the mean estimator variance from 1000 conventional survey simulations by the mean estimator variance of the corresponding 1000 adaptive survey simulations. The efficiencies for the expected nominal encounter rate, $f(0)$ and density estimators for the three population types are given in Table 4.3.

For the clustered populations, adaptive sampling indicated improved density estimate precision with an efficiency of 1.050 for the Clustered and 1.059 for the Highly Clustered populations. As expected, for a CSR population, adaptive sampling efficiency at 0.990 was less than conventional sampling. This is because all animals are randomly located, so increasing the search effort following an observation does not increase the probability of detecting another animal. Thus, with a CSR population, the expected total number of sightings for an adaptive survey is the same as the expected total for a conventional survey. However the sightings in the adaptive survey are then weighted, to account for any adaptive bias, and so there is a decrease in efficiency.

Table 4.3: Efficiencies of adaptive simulation estimates, where efficiency is measured as mean variance of conventional estimator from 1000 simulations divided by mean variance of corresponding adaptive estimator.

Population	Estimator Adaptive Efficiency		
	$\hat{V}\{\hat{E}(e \mid L')\}$	$\hat{V}\{\hat{f}(0)\}$	$\hat{V}(\hat{D})$
CSR	0.959	1.032	0.990
Clustered	0.993	1.265	1.050
Highly Clustered	1.035	1.349	1.059

95% confidence intervals, assuming a normal distribution, for the mean percent relative bias of the encounter rate, $f(0)$ and density estimators are given in Table 4.4. Overall there appears to be no or minimal bias. There is a small negative bias in the $f(0)$ estimate for the clustered adaptive and the conventional highly clustered surveys. This is probably because Akaike's Information Criterion (AIC), which was used to select between the contending models, tended (70% of the time) to select the Fourier series model rather than the (true) half-normal model. There was also a small negative bias for the density estimate of the adaptive clustered survey, presumably largely due to the negative bias in the $f(0)$ estimate.

The mean root mean square errors (RMSE) for the encounter rate, $f(0)$ and density estimators are given in Table 4.5. The adaptive sampling encounter rate estimators do not perform well. However, the improvement in the $f(0)$ estimate outweighs this, leading to an overall improvement in the precision of density estimates.

Table 4.4: Estimated 95% confidence intervals for mean percent relative bias over all 1000 simulations. For each estimator the top confidence interval is for the adaptive survey simulations and the lower one is for the conventional survey simulations.

	Estimator		
Population	$\hat{E}(e \mid L')$	$\hat{f}(0)$	$\hat{D}$
CSR	[-0.91 %, 0.52 %]	[-1.45 %, 0.83 %]	[-1.96 %, 0.07 %]
	[-0.56 %, 0.86 %]	[-1.41 %, 0.16 %]	[-1.54 %, 0.56%]
Clustered	[-1.46 %, 0.32 %]	[-1.84 %, -0.35 %]	[-2.86 %, -0.58 %]
	[-1.32 %, 0.38 %]	[-1.34 %, 0.25 %]	[-2.15 %, 0.15 %]
Highly Clustered	[-2.00 %, 0.19 %]	[-0.67 %, 0.74 %]	[-2.16 %, 0.44 %]
	[-0.50 %, 1.55 %]	[-2.01 %, -0.32 %]	[-1.90 %, 0.78 %]

Table 4.5: Mean root mean square errors for 1000 simulations. For each estimator the top value is the mean for the adaptive simulations and the bottom value the mean for the conventional simulations.

	Mean RMSE		
Population	$\hat{E}(e \mid L')$	$\hat{f}(0)$	$\hat{D}$
CSR	0.00520	0.330	0.0098
	0.00519	0.336	0.0101
Clustered	0.00648	0.321	0.0111
	0.00618	0.341	0.0112
Highly Clustered	0.00804	0.304	0.0127
	0.00752	0.362	0.0131

Table 4.6 shows the coverage of a log-normal 95% confidence interval for the encounter rate, $f(0)$ and density estimators. Values are presented as the percentage of occasions the true value is below or above the estimated confidence interval. Coverage for the $f(0)$ estimates was poor, with the true value being larger than the upper confidence limit for 13% to 16% of the time. Much of this can be explained by the tendency for AIC to select the Fourier series model rather than the half-normal. In general the confidence interval coverage was very similar for the two approaches.

The simulations of heterogeneity in $f(0)$ were analysed using the estimating equations 4.22 and 4.23 for the detection function. These gave improved adaptive density variance estimator efficiencies, of 1.068 for the simulation of increased observer awareness and 1.036 for the simulation of bad weather. However in the case of the bad weather simulation, there was an improvement in the encounter rate efficiency and a decrease in the $f(0)$ efficiency. This was borne out in the mean RMSE's for the density variance estimators. In the increased observer awareness simulation the mean adaptive RMSE (0.0115) improved on the mean conventional estimator RMSE (0.0119), whilst for the bad weather simulation the mean adaptive RMSE (0.129) was larger than the mean conventional RMSE (0.0120). There was a small negative bias in the adaptive density estimate for the increased observer awareness simulation whilst the conventional survey indicated a small positive bias. 95% confidence intervals for the mean percent relative bias were [-3.347, -1.009] and [0.358, 2.786] respectively. The corresponding bad weather simulation confidence intervals were [-8.623, -6.157] and [-6.276, -3.899] for the adaptive and conventional surveys, thus both density estimates demonstrated a negative bias.

Table 4.6: Percentage of occasions the true value is below, above a nominal 95% confidence interval for the estimator, for 1000 simulations. For each estimator the top values are for the adaptive simulations and the bottom values for the conventional simulations.

	Estimator		
Population	$\hat{E}[e \mid L']$	$\hat{f}(0)$	$\hat{D}$
CSR	3.0% , 3.2%	2.9% , 13.2%	2.8% , 5.7%
	3.0% , 2.7%	4.1% , 13.2%	3.4% , 4.9%
Clustered	1.2% , 1.7%	5.3% , 16.2%	1.1% , 3.5%
	0.4% , 1.2%	3.8% , 12.9%	0.8% , 2.3%
Highly Clustered	0.4% , 1.1%	6.3% , 13.1%	0.6% , 2.1%
	0.3% , 0.6%	5.0% , 14.0%	0.4% , 2.5%

4.5 Point Transect Theory

We have so far concentrated on line transects. However the methods are readily extended for point transects. We first need to define some new terminology, so that the methods are more easily transferred.

Assume a point transect survey has been designed to provide even coverage of the survey region. We will refer to the selected sample points as the *nominal points*. Associated with each nominal point is a *location* which may contain both nominal points and additional *adaptive points*.

An amount of additional effort is set aside, above that required to survey the nominal points, to survey adaptive points. A trigger condition is defined, and around any nominal point that meets this condition, additional adaptive points are added based on the effort factor. Thus the effort factor calculation will now need to produce an integer, which will relate to the number of adaptive points to be added to the nominal point.

A number of adaptive patterns are possible, but one option would be to add the adaptive points, equally spaced, on a circle centred on the triggering point. The circle could then be rotated through a random angle between 0 and 2π. Alternatively, a method that is easier to implement in the field, is to cycle through a range of start points on each trigger. So for the first trigger it would start at 0, the next trigger it would start at $\pi/2$, then π, then $3\pi/2$ and finally back to 0. Assuming distances at each point were truncated at radius R, and additional points were added on a circle of radius $2R$, so that the edges of the circular plots touched, then a maximum of 6 points could be added around a triggering point (Figure 4.4). Thus in this case the effort factor calculation would also need to be restricted to an upper value.

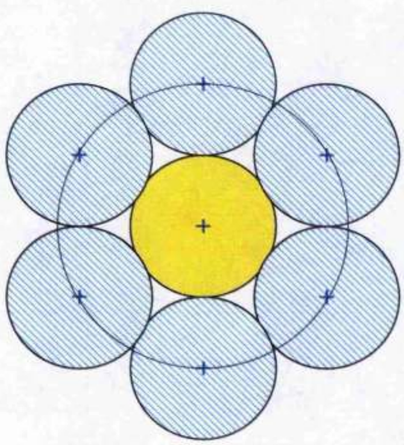


Figure 4.4: Example point transect adaptive pattern, where adaptive points are added in a circle around the triggering point, up to a maximum of 6 additional points. The initial point is shown as a solid (yellow) circle and the potential adaptive points with (blue) cross-hatching. Between 1 and 6 points are equally spaced on the surrounding circle, depending on the effort factor. The surrounding circle of adaptive points is then rotated to a random angle between 0 and 2π.

With this approach the adaptive points themselves do not trigger further adaptive sampling, although variations are possible. In fact many other adaptive patterns are feasible, however for the remainder of this chapter we only consider this one.

4.5.1 Assumptions

The assumptions follow the same format as for the line transect sampling in this chapter. First the following conventional point transect assumptions apply:

i) Probability of detection on the point g(0), is 1.

ii) There is no size bias (the probability of detection is independent of the group size).

iii) There is no responsive movement of animals in advance of detection.

These assumptions can be weakened or removed using similar strategies as for conventional point transect sampling. In addition the following assumptions are made specifically for these adaptive point transect methods:

iv) The expected encounter rate for an adaptive point is the same as the expected encounter rate for the associated nominal point.

v) The expected group size for an observation on an adaptive point is the same as the expected group size for an observation when at the associated nominal point.

vi) Conditional on the location of the actual point, each observation is an independent event. That is, the probability of an observation is only a function of its perpendicular distance from the actual point (although the position of the point itself may depend on past observations).

Notation

Let

k	is the number of points
λ	is the effort factor
n	is the number of animal groups detected
e	is the encounter rate (number of groups detected per point, $e = n/k$
s	is the group size (number of animals observed in the group)

D	is the density (animals per unit area)
$h(0)$	is the derivative of the probability density function $f(r)$ evaluated at r=0

Subscript i is used to refer to the location, i=1..k, and subscript j refers to the points at a location, j=1..m_i, where j=1 always references the nominal point itself. Thus n_{ij} is the number of animal groups detected at the jth point of the ith location. Subscript Y refers to the observation at a point, Y=1..n_{ij}, so that s_{ijY} refers to the group size of the Yth observation of the jth point of the ith location.

We continue the concept of nominal values, where these represent expected values had a conventional point transect survey been carried out, without the adaptive points. As before, nominal effort (now the number of points) is signified with a dash, such as k', the number of nominal points. So, for example, expected number of observations, if only the nominal points had been used, is represented by $E(n \mid k')$. The same approach is also used for both the expected encounter rate and expected group size if only the nominal effort had been used, giving, for example, $E(e_i \mid k_i')$ the expected encounter rate for the ith nominal point, if only that point and none of the associated adaptive points had been sampled.

4.5.2 Effort Factor Calculation

The effort factor for the ith location directly relates to the total number of points sampled at that location (adaptive points and the nominal point), i.e.

$$\lambda_i = k_i \tag{4.40}$$

The effort factor is calculated as a function of the remaining effort available to survey points and the expected number of times the survey will adapt for the remaining survey.

Let

K	be the total number of points to be surveyed
$K_E(t)$	be the total number of adaptive points still to be sampled after time t
$K_R'(t)$	be the number of nominal points still to be surveyed at time t

$K_U(t)$ be the total number of points surveyed by time t

So that

$$K_E(t) = K - K'_R(t) - K_U(t)$$

Let ξ be the expected number of times the survey will adapt for the remainder of the survey (excluding the current increase). Then the increase in effort, following an observation, is given by the excess effort available divided by the expected number of times the effort will increase plus one (for the current increase). So the increase in effort for a nominal point is

$$\lambda_i - 1 = \frac{K_E(t)}{1+\xi}$$

and so the effort factor is given by

$$\lambda_i = RND\left(1 + \frac{K_E(t)}{(1+\xi)}\right) \qquad (4.41)$$

where the function RND(.) rounds the input value to the nearest integer with a maximum value of 7.

As each effort increase can only occur at a nominal point, then ξ can be calculated from an estimate of the trigger rate, the expected number of times the survey will meet the trigger condition per nominal point. If previous survey data is available then a simple estimate of the trigger rate is given by the number of points that met the trigger condition divided by the total number of points surveyed. Let γ be an estimate of the trigger rate, so that

$$\xi = \gamma \cdot K'_R(t) \qquad (4.42)$$

Thus the effort factor is calculated as

$$\lambda_i = RND\left(1 + \frac{(K - K_U(t) - K'_R(t))}{(1 + \gamma \cdot K'_R(t))}\right) \qquad (4.43)$$

4.5.3 Density Estimate

From Buckland *et al* (2001: p55) the density of a population from point transect sampling is given by

$$D = \frac{E(n) \cdot h(0) \cdot E(s)}{2\pi k}$$

where

$E(n)$	is the expected number of animal groups in the sample
$h(0)$	is the derivative of the probability density function $f(r)$ evaluated at r=0
$E(s)$	is the expected group size for the population
k	is the number of points

Replacing the parameters by their estimators then the estimate of density is given by

$$\hat{D} = \frac{\hat{E}(n \mid k') \cdot \hat{h}(0) \cdot \hat{E}(s \mid k')}{2\pi k'} \tag{4.44}$$

Using the delta method (Seber 1982: p5-7), an estimate of the variance of the density estimate is given by

$$\hat{V}(\hat{D}) = \hat{D}^2 \cdot \left[\frac{\hat{V}\{\hat{E}(n \mid k')\}}{\{\hat{E}(n \mid k')\}^2} + \frac{\hat{V}\{\hat{h}(0)\}}{\{\hat{h}(0)\}^2} + \frac{\hat{V}\{\hat{E}(s \mid k')\}}{\{\hat{E}(s \mid k')\}^2} \right] \tag{4.45}$$

4.5.4 Number of Observations

An estimate of the number of observations if only the nominal point had been used for the i^{th} location is given by

$$\hat{E}(n_i \mid k_i') = \frac{\sum_{j=1}^{k_i} n_{ij}}{k_i \cdot \lambda_i} \tag{4.46}$$

with the corresponding survey estimate of

$$\hat{E}(n \mid k') = \sum_{i=1}^{k'} \hat{E}(n_i \mid k_i') \tag{4.47}$$

An estimate of the variance of estimated expected number of observations if only the nominal points had been used is given by

$$\hat{V}\{\hat{E}(n \mid k')\} = \frac{1}{k'-1} \cdot \sum_{i=1}^{k'} \left[\{\hat{E}(n_i \mid k_i') - \hat{E}(n \mid k')/k'\}^2 \right] \tag{4.48}$$

4.5.5 Group Size

The mean observed group size for the j^{th} point of the i^{th} location is

$$\bar{s}_{ij} = \frac{\sum_{Y=1}^{n_{ij}} s_{ijY}}{n_{ij}}$$

and the mean observed group size for the i[th] location is

$$\bar{s}_i = \frac{\sum_{j=1}^{k_i} \bar{s}_{ij}}{\sum_{j=1}^{k_i} n_{ij}} = \frac{\sum_{j=1}^{k_i}\sum_{Y=1}^{n_{ij}} s_{ijY}}{\sum_{j=1}^{k_i} n_{ij}}$$

Assuming there is no size biased detection, and that the expected group size for an adaptive point is the same as the expected group size for the associated nominal point as per assumption v), then following the same argument as for the line transects, an estimate of the expected group size for the i[th] location, if only the nominal point had been sampled, is

$$\hat{E}(s_i \mid k_i') = \bar{s}_i = \frac{\sum_{j=1}^{k_i}\sum_{Y=1}^{n_{ij}} s_{ijY}}{\sum_{j=1}^{k_i} n_{ij}} \tag{4.49}$$

and the estimated expected total number of animals observed for the i[th] location, if only the nominal point had been sampled, is

$$\hat{E}(n_i \mid k_i') \cdot \hat{E}(s_i \mid k_i') = \frac{\sum_{j=1}^{k_i} n_{ij}}{k_i \cdot \lambda_i} \cdot \frac{\sum_{j=1}^{k_i}\sum_{Y=1}^{n_{ij}} s_{ijY}}{\sum_{j=1}^{k_i} n_{ij}} = \frac{\sum_{j=1}^{k_i}\sum_{Y=1}^{n_{ij}} s_{ijY}}{k_i \cdot \lambda_i}$$

Using weighted averages, an estimate of the expected survey group size, if only the nominal point had been sampled, is given by

$$\begin{aligned} \hat{E}(s \mid k') &= \frac{\sum_{i=1}^{k'} \left\{\hat{E}(n_i \mid k_i') \cdot \hat{E}(s_i \mid k_i')\right\}}{\sum_{i=1}^{k'} \hat{E}(n_i \mid k_i')} \\ &= \frac{\sum_{i=1}^{k'} \left\{\hat{E}(n_i \mid k_i') \cdot \hat{E}(s_i \mid k_i')\right\}}{\hat{E}(n \mid k')} \end{aligned} \tag{4.50}$$

An estimate of the variance of the expected group size, if only the nominal points had been used, is

$$\hat{V}\left\{\hat{E}(s \mid k')\right\} = \frac{1}{\hat{E}(n \mid k') \cdot (k'-1)} \sum_{i=1}^{k'} \left[\hat{E}(n_i \mid k'_i) \cdot \left\{\hat{E}(s_i \mid k'_i) - \hat{E}(s \mid k')\right\}^2 \right] \quad (4.51)$$

4.5.6 *h*(0)

As with the line transect estimators, it is assumed there is no correlation between density and $h(0)$ and so observations are pooled across all points (nominal and adaptive) to produce a single estimate using conventional techniques.

4.6 Extensions

4.6.1 Bootstrapping

The bootstrap (Efron and Tibshirani, 1993) is a widely used method for quantifying variance in line and point transect sampling when it is thought that analytic variance estimators do not incorporate all sources of variance.

It is normal to resample transects or points, but this option is not available for the methods of this chapter, as due to the variable effort factor, the transects/points are not independent of one another. If the sightings are assumed to be independent, then it is acceptable to bootstrap observations to obtain variance and mean estimates of $f(0)$ or $h(0)$. If the effort factor is held at a fixed value, such as is done with the experimental harbour porpoise survey in Chapter 5, then the transects/locations can be considered independent (assuming an appropriate survey design) and transects/locations can be re-sampled.

Thus to perform a bootstrap both the population (with appropriate clustering) and the adaptive surveying would need to be simulated for each bootstrap resample. This would require a model of the underlying population distribution to be identified which could then be used to simulate populations. Some work has been done with spatial modelling from line transect data (see for example, Brown and Cowling, 1998; Hedley, 2000; Hedley *et al.*, 1997; and Cowling; 1998), however significant research is still required to identify how best to model the population using the line or point transect survey data, and this avenue is been pursued further within this thesis.

Christman and Pontius (2000) attempt to address bootstrapping a Thompson style simple random sample without replacement adaptive sample by bootstrapping the networks rather than the units. However there does not appear to be a simple correlation to the PB method and so this has not been investigated further.

4.6.2 Coping with Poor Coverage

So far the chapter has concentrated on increasing the effort in areas of high density to increase the sample size. However, the formulae do not require the effort factor to be greater 1. Thus the approach has a second use as it can also be used to compensate for poor survey coverage. Suppose a conventional line transect survey had been correctly designed so as to ensure good coverage, however due to weather and time, or some other such constraints, surveying was only completed for 70% of the trackline in one area. Then the adaptive methods of this chapter could be used, with the effort factor set to 0.7 for transects in the affected area and at 1.0 for all other areas.

Although no simulation has been tried for this approach, effort factors of less than one were recorded in the experimental harbour porpoise survey in Chapter 5. These occurred where the boat went off-effort, for lunch or to investigate a species of interest, and returned to the original transect further along than the position it had originally left. This resulted in one or two legs having an effort factor of less than one, and did not present any issues in the analysis.

4.7 Discussion

Overall the simulation results indicate that conditioning on the effort factors only introduces small bias and that the PB method for line transect sampling offers potential for improving density estimator precision for clustered populations. They also indicate a correlation between the degree of clustering and the adaptive efficiency. Adaptive sampling, in its basic form, will offer no benefit for a population which is not spatially clustered and will in fact be detrimental to efficiency. However, most natural populations will display some clustering, and for populations with high clustering, a benefit is certainly apparent. An indication of clustering is provided by the relative variance, $V(n)/E(n)$ (Cressie, 1993: p590), with the three simulated population types having mean values of 1 for the CSR, 12 for the clustered and 31 for the highly clustered.

The simulations were sensitive to changes in the adaptive pattern. In particular if the adaptive track was too large, so that it frequently stepped outside an animal cluster, this introduced (small) bias into the encounter rate estimate. This is due to violation of assumption (iv), that the expected encounter rate for the adaptive track is the same as the expected encounter rate of the corresponding nominal track. In reality, extra effort is more likely to be triggered when passing near the centre of a cluster, so that adaptive legs may tend to have a slightly lower expected encounter rate than the corresponding nominal legs. It should also be noted that the higher the effort factor the more acute the turn on the zigzags, which may introduce both navigation issues and heterogeneity through problems such as double counting.

The ad hoc approach to handling heterogeneity in $f(0)$ performed reasonably well. For the bad weather simulation, the mean adaptive density estimator RMSE was larger than the mean conventional RMSE, although this can partly be explained by the smaller increase in adaptive observations compared to the other surveys. Poor sighting conditions were simulated for 400 units, meaning that the adaptive survey seldom triggered during this time. Thus the majority of the adaptive triggering occurred during the remaining 900 units of nominal effort, causing larger adaptive zigzags, which for much of the time would then step outside clusters.

Investigation is necessary into how to handle a survey involving multiple species, where similar issues arise to the adaptive methods of earlier chapters. Users could select to trigger on one species only, but then the weights will result in inefficient estimation of density of the other species, unless their areas of high density correspond to those for the trigger species. If the primary species and the secondary species are not spatially correlated by habitat, feeding or other factor, it may be acceptable to treat the sightings for the secondary species as conventional sightings and analyse appropriately.

The effort factor calculation does not adjust for changes in expected encounter rate as the survey progresses, so if there was a density gradient in the survey area, there is the potential for the adaptive algorithm to be inefficient. This may adversely affect precision but bias should be unaffected. To minimise this effect, nominal tracklines should run roughly perpendicular to known density contours. If the gradient was extreme and there were few tracklines, such that there was an excess of additional effort remaining at the end of the survey, there is the potential for the adaptive track to step outside clusters and so induce a small amount of bias.

The efficiency is dependent on appropriately selecting the trigger and stopping function; adaptive pattern; and amount of excess effort available. Thus further work is necessary to estimate the degree of clustering for which adaptive sampling is beneficial, and how to tune the adaptive settings to maximise efficiency. In particular, a number of areas warrant further investigation.

The trigger function is very simple, with effort increased if the number of observations within a section exceeds some value (zero in the simulations). This does not cater for surveys of multiple species, where different trigger functions may be required. The issue is further complicated by the appropriate behaviour of the trigger function during a period of increased effort. Currently, primarily in the interests of acceptable field methods, the effort is not increased further when a detection occurs on a zigzag section. If observations are detected on the last leg of a zigzag, then the effort factor is re-calculated, and a new series of zigzags begins.

The survey returns to nominal effort, following an adaptive trigger, after a fixed number of zigzags. There is potential to develop more sophisticated stopping functions.

The expected encounter rate is fixed at the beginning of the survey, which requires that either an initial estimate (or guess) is available or a pilot survey is carried out. Adjustment of the expected encounter rate using the data that accumulate as the survey progresses may prove useful, particularly when a reliable initial estimate is not available.

The design of the zigzag sections (angle and number of zigzags, and length of section) requires investigation. When each leg in a zigzag is not large relative to the truncation distance W, end and edge effects could be problematic, and field procedures need to be carefully defined to minimize bias (this is discussed further in Chapter 5).

Ideally, simulation will be used to identify suitable parameters, prior to any survey. However as a rule of thumb, the zigzag pattern with 2 or 3 complete cycles performs well and a suitable trigger value could be obtained either from a short pilot survey or by examining previous survey data.

Overall these methods offer a number of advantages over the Thompson-based adaptive methods. Although not design unbiased, the simulations have shown that the bias introduced by conditioning on the effort factors is small. The notation and formulae for the methods are less complex than Thompson's and thus significantly easier to understand and use. This in itself is a significant benefit as Thompson's methods can be extremely involved and it is easy to make a mistake when calculating estimates.

The zigzag adaptive pattern allows surveying to be continuous for a line transect survey , and so removes the need for off-effort. Depending on the survey this may offer a distinct advantage. If observers can move easily between transects, but the transects themselves are surveyed slowly, for example crawling on hands and knees looking for deer dung, then the off-effort will be a negligible factor and may be seen as a welcome break. However if the spacing to adaptive transects is large and the speed of travel is comparable to the survey speed, then resources may not be used optimally and potential surveying time maybe wasted.

The harbour porpoise survey in Chapter 5 pointed to poor performance of the group size estimator, although the reasons for this are not fully understood. Ideally the simulations would have estimated group size, instead of modelling observations as single animals, although to provide a useful comparison the simulations would have also needed to model the bias. Use of the more sophisticated estimator utilising covariates in the detection function estimate may to go some way to addressing these issues. However tests also need developing to detect heterogeneity in $f(0)$.

Applying the methods to point transect surveys needs further work, and in its current form may offer less benefit than for line transects. The approach currently requires the effort factor to be a multiple of the effort to survey a single point, and thus it is difficult to tune parameters to a population's spatial distribution and the mechanism is likely to be highly sensitive to slight changes in the underlying model. It may be preferable to consider adding adaptive points to groups of points, located within a common area, rather than to individual points. The advantage of this is that effort could be increased in smaller increments, and thus be less susceptible to changes in the spatial clustering. The basic concepts remain the same, and using the terminology of the section, you would still have locations, but now a location would consist of more than one initial point. In making this change the estimators will need to be slightly modified, bringing them closer to the line transect formulae. This revised approach would also open up the options for many other adaptive patterns.

So far, the focus of the thesis has been the improvement of estimator precision by increasing the sample size. However the PB method also provides the ability to improve survey coverage through the variable effort factor. It is likely that this will prove the main benefit of the approach. Because the effort factor is a function of whether the survey is ahead of or behind schedule, the method can accommodate some loss of effort due to poor conditions. Effort simply resumes when conditions improve, and the effort factor in adaptive legs is reduced accordingly. Additionally, the methods can also accommodate an effort factor of less than one. Thus areas with incomplete surveying, due say to bad weather, can still be included in the analysis without biasing the abundance estimate.

In the next chapter we report on the application of this adaptive approach to an experimental line transect survey and discuss the field methods.

Chapter 5

Harbour Porpoise Survey

5.1 Introduction

In this chapter we apply the adaptive PB method, developed in Chapter 4, to experimental line transect surveys of harbour porpoise (*Phocoena phocoena*). The surveys were performed between the 6th and 28th August 1996 in the Gulf of Maine/Bay of Fundy area by the Northeast Fisheries Science Center (NEFSC), Woods Hole, Massachusetts, USA. The purpose was twofold: to understand the practical aspects of adaptive sampling, such as observer methods; and to provide a real life analytical comparison of the method with conventional line transect sampling. The results were promising, showing both that the approach was practical and that precision was increased for the adaptive estimates.

Comparative conventional and adaptive line transect surveys were performed over the same nominal transects to provide a direct comparison of the two methods. Harbour porpoise were considered a good species to test the approach on, as they have a relatively high abundance in the region. This provided sufficient observations for both types of survey to be carried out over the same transects within a single day. A number of days were lost or surveys aborted due to bad weather, and in total seven comparative conventional/adaptive survey pairs were conducted across four locations in the area. The separate surveys were pooled and analysed as a single large survey.

Surveying was performed from the *R/V Abel-J*, a ship used in previous harbour porpoise line transect surveys. The Abel-J is 32m long with a 6m draft. The vessel has been extensively dampened, includes a number of observation platforms, and is highly suited to marine mammal surveys.

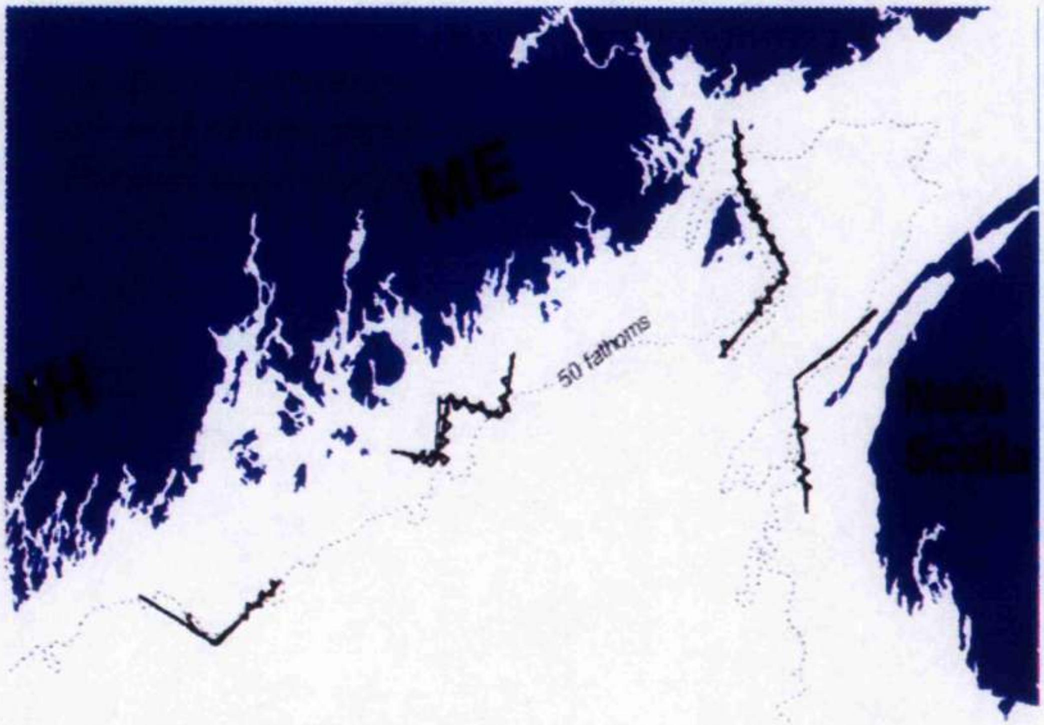


Figure 5.1: Track lines surveyed during conventional and adaptive line transect surveys of harbour porpoise.

The chapter describes the initial design of the survey and the selection of the adaptive parameters as well as the survey procedures. It goes on to describe the analysis of the data, presents the survey results and finishes with a discussion of the experiment conclusions.

Throughout this chapter, school size is used to refer to the number of animals in a single observation (the group size).

5.2 Survey Methods

5.2.1 Survey Design

To simplify the approach and navigation for this first trial, it was agreed to fix the effort factor to a constant value rather than vary it dependent on the remaining available effort. Each conventional and adaptive survey pair followed the same nominal transects, and so the adaptive surveys used a greater total effort than the conventional surveys. Hence adjustments were required in the analysis to account for this. The surveying was conducted in areas of expected high porpoise densities with the sole purpose of testing the adaptive theory and methods by comparing with conventional line transect sampling. Thus the survey data should not be used to produce an actual density estimate for the Gulf of Maine/Bay of Fundy region.

Selecting Survey Parameters

In preparation for the survey the RATS simulation program (Appendix A) was used to identify an appropriate adaptive pattern, effort factor and trigger value. This was a two stage process; first the population simulation parameters were adjusted in an attempt to match abundance estimates from previous surveys of the area. Then using populations, generated with these parameters, repeated simulations were run with varying surveying parameters to identify a suitable configuration for the survey. There was little preparation time prior to the survey, and RATS has limited population simulation capabilities, thus in the interest of expedience a fairly basic approach was followed.

To select suitable population parameters, data from previous surveys of the area (Smith *et al.*, 1993, Palka, 1995 and Palka, 1996) were plotted, typically for a day at a time as this produced a reasonable size plot. Conventional line transect surveys were then simulated using the half-normal detection function, with parameter σ = 300 metres, and a truncation half-width of 500 metres, so that the ESW (effective strip width) was approximately 340 metres. This was a passable approximation to previous survey results. Survey parameters were adjusted by eye until detection patterns were typically similar to the plotted detections from previous surveys, with a target population density of approximately 0.7 schools per square kilometre.

Once identified the population parameters were used for multiple simulation runs to test two adaptive patterns, a zigzag and a hounds-tooth (Figure 5.2).

For each pattern a single cycle consisted of a short leg up, a long leg down and then a short leg back up. For the simulation, the first leg of the adaptive pattern went in the same direction each time, whilst for a real survey the pattern should ideally reverse on each non-contiguous trigger. For example if it went up on the first leg for the previous trigger, it should go down for the first leg on the next trigger. If there were sufficient sightings to trigger the adaptive condition on the last leg of the pattern, then rather than revert to a straight line, the adaptive pattern was repeated.

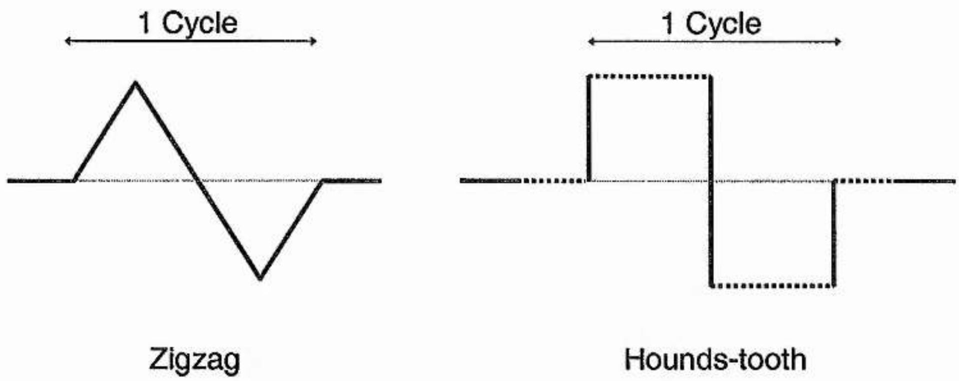


Figure 5.2: Zigzag and hounds-tooth adaptive patterns, each shown for a single cycle. The off-effort travel is shown as a dotted line.

The zigzag used continuous effort with no break in surveying when the direction changed, as the angle of turn was typically felt low enough for surveying still to be feasible. This aspect is considered further in the Field Procedures part of this chapter.

The hounds-tooth pattern could more accurately be thought of as a series of transects perpendicular to the nominal transect, and involved a degree of off-effort. With this adaptive pattern the perpendicular transects were spaced twice the truncation half-width distance apart and there was no sampling on the interconnecting transects, which approximated for the off-effort lost in turning through 180 degrees (Figure 5.3). In addition an off-effort buffer zone of the same length as the truncation half-width is also added at the start and end of the complete adaptive pattern, again to account for the turn. This buffer zone was only added at the start and end of the whole pattern and not between each cycle, so the more cycles the lower the proportion of off-effort.

The additional sections of off-effort also prevented re-sighting of animals in overlapping transects, otherwise the detection function would not be accurately modelled. This is because on a turn, the location of observations would already be known for the overlapping part of the transect and the porpoises are unlikely to have moved very far within such a short space of time.

The effort factor for the hounds-tooth pattern was calculated as the length of the perpendicular transects divided by the nominal length of the pattern, and so did not include the off-effort sections. The hounds-tooth pattern requires the perpendicular transects to be reasonably long to provide an opportunity for animals to be detected and also to reduce the proportion of off-effort.

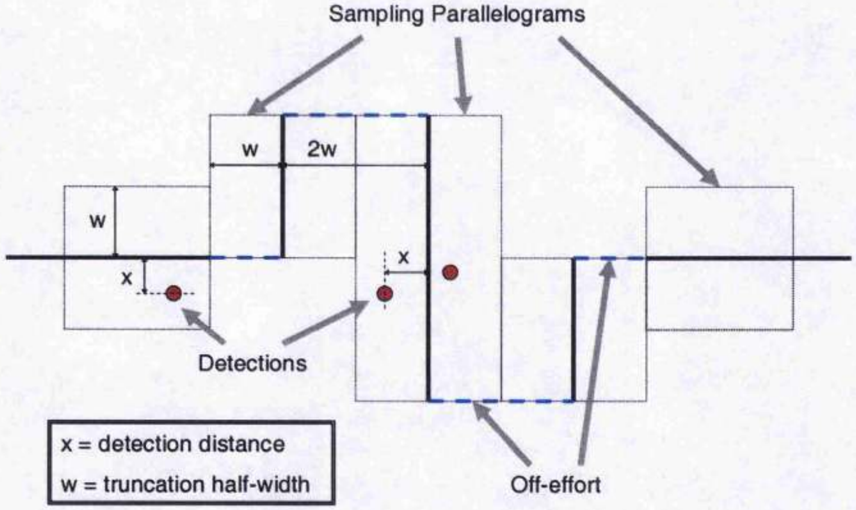


Figure 5.3: Surveyed sub-transects for hounds-tooth adaptive pattern.

Initial assessment of the two patterns was made by performing runs of 1000 simulations, comparing first the adaptive zigzag with a conventional survey and then the adaptive hounds-tooth with a conventional survey, where the populations for each run were generated using the estimated population parameters. The adaptive pattern was set to only repeat for a single cycle, and the effort factor for the zigzag pattern was set to 3 whilst the effort factor for the hounds-tooth pattern was set to 5. The efficiency of the density estimator was measured as variance of the conventional estimate divided by the variance of the adaptive estimate. The mean zigzag density estimator efficiency was 1.17 and the mean hounds-tooth efficiency 1.14, although these estimators had not been adjusted for the differences in total effort. The total effort for the conventional survey was set to 800 kilometres, whilst the nominal effort for the adaptive surveys was set to 600 kilometres. The mean total effort used by the zigzag pattern was 779 kilometres (i.e. less than the total used for the conventional survey), whilst the mean total effort for the hounds-tooth pattern was 838 kilometres with 60 kilometres of off-effort. Thus the hounds-tooth efficiency is reduced further if adjusted to a comparable total effort. A single cycle with the hounds-tooth pattern gives the highest ratio of off-effort to surveyed transect. Thus increasing the number of cycles would have reduced the amount of lost off-effort, although at the risk of stepping outside the cluster with the adaptive pattern and therefore decreasing the adaptive efficiency.

Further simulations were carried out varying the effort factor and trigger values to investigate behaviour. Example simulations for a zigzag and a hounds-tooth pattern are shown in Figure 5.4. Experimentation indicated that for these populations, with a single adaptive pattern cycle, the optimum fixed effort factor for the adaptive zigzag pattern was a value of 2 to 3 and for the hounds-tooth 4 to 5. However due to the makeshift nature of the population matching, it would be unrealistic to draw general conclusions from this; rather this is an indication of the general parameters that may suit this particular survey.

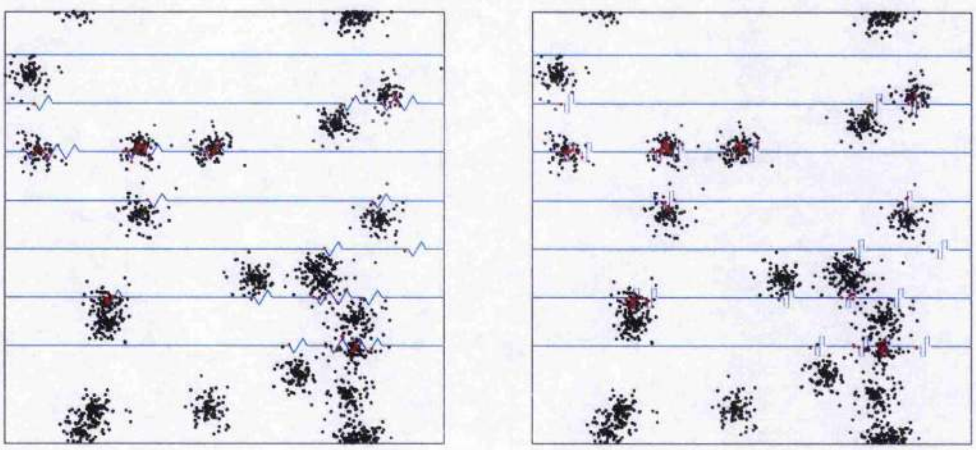


Figure 5.4: Example simulations using a zigzag pattern on the left and a hounds-tooth pattern on the right. In each case the adaptive pattern repeated for a single cycle following a trigger. The trigger was 1 observation in the previous unit. Black dots represent the (undetected) animals in the population, whilst detected animals are shown as a red dot. Transects are shown in dark blue.

The simulations showed that the hounds-tooth pattern had to be carefully tuned with respect to its length and height to improve density estimator precision over a conventional survey. Whilst the zigzag pattern also required careful tuning, it was more robust to changes in the population. In addition as the zigzag pattern did not involve off-effort travel, it had an inherent efficiency advantage.

As a result of this, and also in consideration of the ease of navigation, the single cycle zigzag pattern was chosen with a fixed effort factor of 2. These values were chosen not just as the most efficient figures for a particular population but also for their robustness with varying populations. Examination of the previous survey data indicated that a realistic trigger was an encounter rate of approximately 0.5 observations per kilometre in low density areas and 1 observation per kilometre in high density areas.

To allow conventional and adaptive line transect sampling techniques to be compared under as similar conditions as possible, each pair of small surveys was typically conducted during the same day in the same region. So, for example, when the weather was predicted good for the entire day, then in the morning one approach would be conducted and in the afternoon the ship would return on the same transect lines using the other technique.

5.2.2 Field Procedures

Standard one-team line transect sighting protocols were used. The observers worked on a single observation deck, 14m above the sea surface and 6m from the bow. A team of 5 observers collected sighting data, with 3 observers working at a time, one observer concentrating to the left, one ahead and one to the right. Observers cycled positions, at 30-minute intervals, starting on the left, moving through centre to right, and then to a rest period. Lots were drawn each morning to select the initial start positions.

Observation was with the naked eye, though 7x50 binoculars were available to confirm school size and species if necessary. Observers were trained to pay particular attention to the area on the outside of each turn, where there is greater area to cover than on the inside. This is highlighted by the shaded triangles in Figure 5.5, in which the strip survey is conceptually shown as a series of parallelograms; ideally survey design should keep these triangles small relative to the survey strip, to reduce the turning effect. During the turns, on the adaptive zigzags, the central observer concentrated more on the outside of the turn, to compensate for the greater area being covered by the outside observer. The observer on the inside of the turn increased their prime area of scanning to include more of the central zone, as they then had less area to cover.

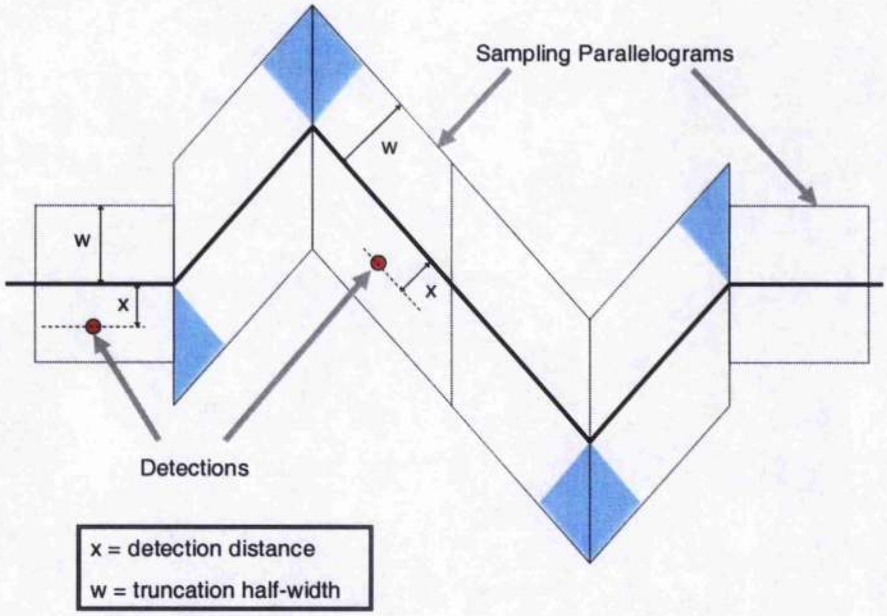


Figure 5.5: Zigzag adaptive pattern. The area is sampled with a series of parallelograms. Additional observer effort, signified by the shaded areas, is required on the outside of turns, whilst reduced effort is required on the inside of turns.

Harbour porpoise are small (adults are approximately 1.5 metres in length) and dark, and typically only the fin and a small proportion of the back is visible to observers. Thus surveys require a good sea state, with calm, flat conditions. Surveying was performed while travelling at 9-10 knots and was generally only conducted in Beaufort sea state 3 or less, although during some of the adaptive surveys the Beaufort sea state increased above 3.

For the adaptive surveys the trigger to increase the effort (zigzag) was based on the number of animals seen within the previous 15 minutes (approximately 4 kilometres). The trigger value was held constant for the individual surveys, but across all the surveys two values were used, either 2 or 4 animals. The higher value of 4 was used where the encounter rates were consistently high and a lower value could have led to repeated adapting.

As an example, consider the use of a trigger value of 4. If a sighting was made, and there had been 3 previous sightings within the last 15 minutes, then the observers would call through to the bridge and the ship would change course to start adapting. There was typically a 1 or 2 minute lapse between the sighting and the change in direction. If the ship was following an adaptive track, then the ship would continue zigzagging if the number of sightings on the last leg of the zigzag was equal or greater than half the trigger value in use.

It was also decided that an adaptive section would not be started if the section could not be completed before the end of the transect.

Data Recording

Each observer used a hand-held computer, a PINGLE (Garret-Logan and Smith, 1997), equipped with an electronic pen for data entry. All species observed were recorded, though triggering for adapting effort was only based on harbour porpoise. Immediately on sighting an animal group (school), the observer clicked a button on the PINGLE to record the time accurately. This enabled the matching of sightings with the transect data, which were recorded separately. If the sighting was invalid, the observer could then cancel the recording. Otherwise the observer proceeded to input the sighting data. These included the following details: time; radial distance to animal, or centre of the school of animals (estimated by eye); the angle between the sighting and the trackline (an angle board was provided for each observer); species; best, high and low estimates for the school size; travel direction of animal/school; number of calves in the school; sighting cue; and animal/school behaviour.

In addition, effort and environmental data were recorded using two other computers managed by a sixth observer on the bridge. The PINGLE clocks and the two bridge computer clocks were synchronised within a few seconds to ensure accuracy of the sighting data relative to the transect data.

One of the computers was a GPS logger recording the following every minute: time; latitude; longitude; ship's bearing and speed; wind speed and direction; Beaufort wind force; bottom depth; surface water temperature; and drift and set from the track line.

The sixth observer used the other computer to record transect-related information, at changes of weather, observer shift, ship's speed or ship's direction. The data recorded included: time; position of each observer; and weather conditions (swell direction and height, Beaufort sea state, presence of rain or fog, percentage of cloud coverage, visibility to the horizon, vertical and horizontal position of the sun, and glare width and strength).

5.3 Analysis

To compare the two methods, point estimates and coefficients of variation (*cv*s) of $f(0)$, expected school size, encounter rate, density of schools and density of individual porpoises were estimated. For simplicity and to increase sample size, transects using the same sampling method from the individual surveys were pooled, to create a single conventional and a single adaptive survey.

5.3.1 Data Preparation

Although the survey ran for 23 days, with travel to the survey area and days lost due to bad weather, there were only 7 pairs of conventional and adaptive surveys to compare. Prior to analysis the data were prepared as follows. First only the transects and sightings for the selected 7 survey pairs were extracted, and only sightings of harbour porpoise were retained. Then any sightings with incomplete data (missing school size, radial distance or angle) were removed. There were two extremely large school sightings in the adaptive data of 75 and 35 animals, whilst the largest school in the line transect data was 15 animals. In general school sizes were similar for the two survey types with the exception of these two very large outliers. It was therefore decided to remove these large school sightings which would otherwise have had an undue influence on the variance.

All of the conventional line transect surveying was performed in Beaufort 3 or less, whilst for some of the adaptive sampling, although started at 3 or lower, the Beaufort rose above this. Due to their size and colour, sighting of harbour porpoise is extremely difficult in Beaufort 4 or above and thus any transect legs performed at this level, and their associated sightings, were removed.

It should also be noted that some of the transect sub-legs were not contiguous. This is due to short breaks off-effort, either to inspect specific animals or for lunch, where the transect was not rejoined at exactly the same location as surveying had previously finished. This meant that on a couple of the adaptive sections the effort factor was actually less than 1. This is because the nominal effort was calculated as the distance from the start latitude and longitude to the end latitude and longitude; and so was potentially less that the actual survey effort if there had been no adaptive zigzags. This should not present an issue for the methods, and in fact it is one of the benefits that the estimators are able to accommodate some incomplete coverage of an area.

5.3.2 Comparison of Results

Comparison of the results is not straightforward: the two survey types used differing total effort; and as this was a field experiment the true parameters, for the two survey types, may not have had equal means. For the PB method simulation in Chapter 4, the efficiency was measured by the ratio of the conventional and adaptive estimator variances, which assumes the means are equal. Here the results are compared by dividing the *cv* of the conventional estimate by the *cv* of the adaptive estimate, and this measure is referred to as the *adaptive improvement.*

For the adaptive surveys the total effort was 1.204 times greater than the total effort for the conventional survey. To account for this, the conventional density estimate *cv*s were adjusted using the approximation (Buckland et al., 2001: p243),

$$L_A = L_C \cdot \left\{cv\left(\hat{D}_C\right)\right\}^2 / \left\{cv_R\left(\hat{D}\right)\right\}^2$$

where

L_C	is the total effort for the conventional survey
$cv\left(\hat{D}_C\right)$	is the *cv* of the density estimate for the conventional survey
L_A	is the total effort for the adaptive survey
$cv_R\left(\hat{D}\right)$	is an estimate of the *cv* of the density estimate for the conventional survey, had it used the same total nominal effort as the adaptive survey.

Thus an estimate of the *cv* of the density estimate for the conventional survey (had it us the same total nominal effort as the adaptive survey) is given by

$$\hat{cv}_R(\hat{D}) = \sqrt{L_C / L_A} \cdot cv(\hat{D}_C) \tag{5.1}$$

The same method was used to revise the conventional *cv*s of the other survey estimates, $\hat{f}(0)$, $\hat{E}(n \mid L')$ and $\hat{E}(s \mid L')$.

In addition, a randomisation version of Levene's test (Levene, 1960; Manly, 1997) was used to test whether the conventional and adaptive variances were equal, and a randomisation version of a one-way ANOVA was then used to test if the means varied between the two survey types. The advantage of Levene's test is that it does not assume the sample means from the two survey types are equal. The randomisation version of the test compares the estimate with a distribution that is generated by randomly reordering the data. In comparison with many other tests, the test has been shown to be relatively powerful and robust to non-normality and differing sample means. In both cases the computer package RT (Manly, 1996) was used to perform the tests.

5.3.3 Parameter Estimation

Both the conventional and the adaptive sightings were analysed with DISTANCE 2.2 (Laake *et al.*, 1994) to produce adaptive and conventional survey $f(0)$ estimates. The truncation half-width w, was set to 700 metres. The detection function model was selected from the Fourier, half-normal with Hermite adjustments and hazard-rate with no adjustments models. Model selection was performed using the AIC (Akaike Information Criterion) values and the hazard-rate model was selected in each case

For the adaptive survey, the encounter rate, school size, school density and individual harbour porpoise densities were estimated using equations 4.9 to 4.19, from Chapter 4, coded in S-Plus version 4.5 for Windows (Statistical Sciences Inc., 1997). For the conventional survey, DISTANCE 2.2 was used to produce the estimates. Note that for a conventional survey, the effort factor is 1, and the PB method equations then simplify to the conventional estimators used in DISTANCE.

Analysis of the conventional sightings, using DISTANCE, showed signs of school size bias when using a truncation half-width of 700 metres. The same was also true of the adaptive sighting data. Reducing the truncation half-width to 400 metres yielded a non-significant relationship between g(y) and the natural log of the school size. Thus for both the conventional and adaptive survey school size estimates only, a truncation half-width of 400 metres was used.

As harbour porpoise can be difficult to sight in even low Beaufort, potential biases due to weather conditions were accounted for by stratifying the data by Beaufort sea state and estimating density for each stratum. The data were stratified into two groups, Beaufort sea states 0 and 1, and Beaufort sea states 2 and 3.

Other methods to account for school size and weather condition biases that were investigated included the bivariate hazard rate model (Palka, 1993) and regression method (Buckland et al., 2001). These other methods produced similar results and have not been included here.

Overall parameter estimates were produced by averaging the stratum estimates, weighted by the proportion of effort in each stratum. The variance and *cv* of all parameters were estimated by bootstrapping (Efron and Tibshirani, 1993) 500 times, resampling transects, with the restriction that the total effort should be within 5% of the base sample. In each case the detection function was set to the hazard-rate model. Normally, with the PB method, it would not be valid to bootstrap transects; as the total effort available at the start of each transect is dependent on the sightings made on previous transects and thus they are not independent. However, in this case the effort factor was fixed and so it was appropriate for this survey.

5.4 Results

The seven comparative survey pairs (labelled A through F) were carried out in four different regions (R1 through R4) and are summarised in Table 5.1. The surveys were performed in areas of expected high porpoise population in order to get sufficient sightings for both the conventional and the adaptive approaches to estimate $f(0)$.

Table 5.1: Summary of survey descriptions for the individual surveys. Surveys are grouped in matching pairs. The trigger value is not used for a conventional survey and so is marked n/a (not applicable).

Pair	Date	Region	Type	No. of Transects	Trigger
A	11 Aug	R1	Conventional	4	n/a
A	11 Aug	R1	Adaptive	4	2
B	13 Aug	R1	Conventional	4	n/a
B	13 Aug	R1	Adaptive	4	4
C	15 Aug	R2	Conventional	3	n/a
C	14-15 Aug	R2	Adaptive	3	4
D	20 Aug	R3	Conventional	2	n/a
D	20 Aug	R3	Adaptive	2	4
E	25 Aug	R4	Conventional	2	n/a
E	25 Aug	R4	Adaptive	1	2
F	26 Aug	R2	Conventional	2	n/a
F	26 Aug	R2	Adaptive	2	2
G	14 Aug	R2	Conventional	3	n/a
G	24 Aug	R2	Adaptive	3	2

Typically an adaptive and a conventional survey were performed on the same transects on the same day. However adaptive survey C was started one evening and was then completed the next day and pair G is made up of surveys performed on differing days. Finally with legs removed due to Beaufort being greater than 3, the adaptive survey E lost a complete transect, and so the overall adaptive survey had a smaller nominal effort than the overall conventional survey. In total there were 19 adaptive transects and 20 conventional transects in the data analysed.

The data from the seven comparative surveys were pooled to create one conventional survey and one adaptive survey. Total effort for the conventional survey was 233.9 nautical miles with 313 sightings, before truncation, and total effort for the adaptive survey was 281.7 nautical miles (nominal effort 183.1 nautical miles) with 551 sightings, before truncation. Sightings were truncated at 700m giving 303 conventional sightings and 523 adaptive sightings (312.2 nominal sightings). The effort data are summarised in Table 5.2 and the truncated sightings data in Table 5.3.

As mentioned in section 5.3.3, the data were stratified by Beaufort sea state into two strata. For the conventional survey 29% of the effort was at Beaufort 0 or 1 and 71% at Beaufort 2 or 3, whilst for the adaptive survey the figures were 42% and 58% respectively (Table 5.4).

Table 5.2: Summary of survey effort and sightings prior to truncation for the conventional and adaptive surveys.

Survey	Total Effort (km)	Nominal Effort (km)	Total Sightings (no truncation)
Conventional	233.9	233.9	313
Adaptive	281.7	183.1	551

Table 5.3: Summary of number of sightings within 400m and 700m of the track line for the conventional and adaptive surveys.

	Truncation 400m	Truncation 700m	
Survey	Number of Sightings	Number of Sightings	Nominal Number of Sightings
Conventional	252	303	303
Adaptive	409	523	312

Table 5.4: Percentage of effort (track length) surveyed in Beaufort sea states 0 to 3, after tremoving transects with higher Beaufort sea state.

	Beaufort Sea State			
Survey	0	1	2	3
Conventional	0	29	50	21
Adaptive	4	38	42	16

A summary of the estimates for the pooled surveys is given in Table 5.5. The conventional estimates have had their *cv*s adjusted to compensate for the greater total effort used for the pooled adaptive survey. Two density estimates are provided, one corresponding to individual porpoises ($\hat{D}$) and one to porpoise schools ($\hat{D}_S$).

Table 5.5: Summary of analysis estimates for harbour porpoise surveys. %*cv*s are shown in brackets. *cv*s for the conventional survey have been adjusted to compensate for the greater total effort used by the adaptive survey. The adaptive improvement is measured as the conventional estimate *cv* divided by the associated adaptive estimate *cv*.

Survey	$\hat{E}(e \mid L')$ (Schools per Nm)	$\hat{E}(s \mid L')$	$\hat{f}(0)$ (Nm^{-1})	$\hat{D}$ (Porpoises per Nm^2)	$\hat{D}s$ (Schools per Nm^2)
Conventional	1.30 (12.2)	2.21 (8.9)	8.24 (14.9)	11.24 (21.5)	5.05 (19.7)
Adaptive	1.71 (10.2)	2.43 (9.1)	5.24 (14.1)	11.07 (19.8)	4.71 (17.4)
Adaptive Improvement	1.20	0.97	1.06	1.08	1.14

The encounter rate estimate, $E(n \mid L')$, for the pooled adaptive survey at 1.71 schools per Nautical mile (%cv=10.2) was significantly greater than that from the pooled conventional survey at 1.3 (%*cv*=12.2). The school size estimate, $E(s \mid L')$, was also greater for the pooled adaptive survey (2.43, %*cv*=9.1) than the pooled conventional survey (2.21, %*cv*=8.9). Estimates of $f(0)$, in Nautical miles^{-1}, for the adaptive and conventional line transect data were 5.24 (%*cv*=14.1) and 8.24 (%*cv*=14.9), respectively. Resultant density estimates and *cv*s from the pooled adaptive survey were slightly less than that from the pooled conventional survey.

By the randomised Levene's tests all parameters, except the encounter rates, had a significant difference in the variances for the two survey types. In addition all parameter point estimates were significantly different, with the exception of the school density estimate, $\hat{D}_S$. The variance of the school density estimate was significantly less for the adaptive survey than the conventional survey.

The adaptive PB method showed an improvement in the density estimates above the conventional methods, where the improvement is measured by dividing the *cv* of the conventional estimate by the *cv* of the adaptive estimate and from the results of Levene's test. The results indicate a 14% improvement in the adaptive school density variance estimate and an 8% improvement in the adaptive individual porpoise variance estimate. The improved efficiency was significant. The lower improvement for the individual porpoise estimate is believed to be due to poor performance of the adaptive school size estimate.

5.5 Discussion

Applying the adaptive survey methods was relatively straightforward. This was aided in that the effort increase was fixed for these surveys and so no calculation was required prior to adapting. The surveys were also carried out in open sea and so land avoidance was not typically a problem.

It was important that the observers were notified as soon as the ship was starting to turn so that they could concentrate more on the outside of the turn. Sightings on a turn had to be treated carefully and observers needed to check that the angle board was correctly lined up immediately following a sighting. In addition it was important to check that a sighting from the previous transect was not double counted in the current transect following a turn.

Care was required when counting sightings on the last leg of an adaptive zigzag, where these sightings were used as a trigger for continued adapting. One approach to ensure the sightings are actually on the last leg is to wait until a straight line is apparent in the ship's wake before counting such sightings towards an adaptive trigger. Alternatively wait a fixed time (say 1 minute) following an announcement from the bridge that the boat is turning.

The experience of the survey shows that the adaptive procedures are viable for observers. The key requirements for success are: good communication between the bridge and the observers; and thorough observer training, ensuring observers have an appreciation of the method and any potential to introduce errors. The inclusion of a variable effort factor in an adaptive survey would involve some basic calculations to be performed at the bridge before turning. However this calculation could easily be simplified using a simple computer application. It is important that the observers appreciate the potential issues involved during a turn on an adaptive survey. As long as they are adequately briefed then it is relatively straightforward for observers to compensate for the turning effects.

To use the adaptive method, an adaptive trigger must be chosen *a priori*. This was possible for this survey because previous harbour porpoise line transect surveys had been conducted in the area. If this had not been the case, then it may be possible to estimate a trigger using: survey data from other areas; biological knowledge of the animals; or a short pilot survey. If the effort factor is fixed, as in this case, then it would be acceptable to change the trigger part way through a survey, although not midpoint in a transect. For this experimental survey, two trigger values were used (2 and 4 schools per 15 minutes). From preliminary investigations, there did not appear to be a significant difference in results due to the different values.

The encounter rate increased for the adaptive survey, which may be partly due to a higher proportion of the conventional survey, relative to the adaptive survey, being carried out in Beaufort 2 or 3, when sighting would be more difficult.

The adaptive method resulted in larger, more variable estimates of average school size. This could be due to a number of factors. First, there were several large schools in the adaptive survey data (in addition to the two very large schools discarded during data preparation). In the conventional data the largest school was 15 porpoises, while there were four larger schools (with 16, 16, 18 and 20 porpoises) in the adaptive data, representing 1% of the total sightings used for the school size estimate. In addition, reducing the truncation half-width from 700 metres (where school size bias was detected) to 400 metres (where school size bias was not detected) probably did not completely account for the bias. Finally, the adaptive method itself will be biased if there is a correlation between school size and encounter rate. The simulations of the method have so far not investigated the school size estimator.

The density estimates from the adaptive survey were lower than those from the conventional survey, which appears to be partly due to a low estimate of $f(0)$. It is not known whether this was a result of size bias or other heterogeneities that were not completely compensated for; a bias in the methodology; or chance, and there were actually less animals in the area when the adaptive surveys were conducted. The simulation of the PB method in section 4.4 indicated that density estimates were not significantly biased by the method. It was assumed that $f(0)$ was not biased by the increased effort in the adaptive surveys. This may not be the case if there was heterogeneity in sighting conditions and the stratification by Beaufort sea state did not completely account for this heterogeneity. Further investigation is required, and it is likely that the $f(0)$ estimator including covariates, developed in section 4.3.3, will prove a more robust estimator, allowing more of the sighting factors to be taken into consideration.

As already mentioned, the degree of spatial clustering, the effort factor calculation and the trigger function are key factors in the efficiency of adaptive surveys. Additional work is necessary to identify at what level of clustering an adaptive survey is worthwhile, which leads to the more complex task of identifying the spatial clustering of a population. Methods are needed to select an appropriate adaptive pattern and trigger function, which are both highly dependent on the population clustering.

In summary the methods proved both practical and beneficial to apply to the harbour porpoise survey with an 8% improvement in the porpoise density estimate *cv* over the conventional survey, a difference that was statistically significant. The survey was fortunate as there was sufficient information available to allow appropriate adaptive parameters to be reasonably estimated prior to commencing. The experiment did not however exploit the method's potential to enable a survey to complete to a fixed time and cost. It is expected that this may well prove the most significant benefit as it allows additional effort assigned to a survey, to account for bad weather for example, to be efficiently allocated throughout the survey.

Chapter 6

Discussion

6.1 Introduction

We started this thesis with the premise that adaptive sampling has the potential to improve the precision of distance sampling abundance estimates for spatially clustered populations. Simulation has shown this to be true in specific cases, but with a number of provisos.

For a population that exhibits complete spatial randomness (CSR) then there will be no benefit and in fact there is likely to be a loss in precision (as demonstrated by the PB method simulations). This is because any adaptive trigger, for say an increased encounter rate, is pure chance; if animals are located at random, there is not an expected increase in observations in the surrounding area.

As the degree of spatial clustering increases then the efficiency of the adaptive methods is likely to increase. This is again borne out by the PB method simulations where the density estimate efficiency was 1.05 for the clustered population and 1.06 for the highly clustered population. Brown & Manly (1998) also suggested adaptive sampling worked well for highly clustered populations, and that there was a level of spatial clustering below which the adaptive method ceased to be more efficient than conventional methods. Further research is required to identify the appropriate level of clustering at which the adaptive procedure is beneficial. An indication is provided by the relative variance (Cressie, 1993: p590) where for the PB method line transect simulations the mean values were 12 and 31 for the Clustered and Highly Clustered population respectively.

The requirement to have previous survey data, or to perform a pilot survey so that appropriate survey configuration can be investigated using simulation, is perhaps unrealistic, and may limit the approach to populations with a well understood spatial distribution from repeated surveying over time.

In reality the gains in adaptive precision are small, particularly when considering the additional survey complexity in theory, analysis and survey procedures. The real benefit may lie with the PB method, which as well as the potential to improve precision has the ability to complete a survey to a fixed time and cost. This maximises the use of survey resources, particularly important if expensive resources such as a ship or aircraft are being used. The PB method also has the unexplored capability to use survey data with an effort factor of less than one. This may prove a useful way to cater for incomplete survey coverage of a portion of the survey region, due to say bad weather. Thus the incomplete area is analysed using an effort factor based on the proportion of completed sampling in the region, allowing these data to be included in the estimate without biasing the abundance estimation.

6.2 Survey Design Considerations

The efficiency of adaptive estimators and particularly Thompson's Hansen-Hurwitz and Horvitz-Thompson-based estimators have been investigated (see for example, Thompson, 1996; Christman, 1997; Brown, 1999) with the conclusion that the Horvitz-Thompson-based estimator can offer the lower variance under the situations explored. Therefore it is suggested that the adaptive point and line transect estimators based on these estimators should typically be chosen in preference to the Hansen-Hurwitz estimators.

Experience with adaptive survey simulations suggested that the survey parameters, such as the trigger condition, the available excess effort, and the adaptive pattern, have to be carefully tuned to suit the underlying population distribution. For example if the adaptive pattern was large so that the increased sample area was much bigger than the actual cluster of animals, then the methods may fail to get a worthwhile increase in observations. Figure 6.1 shows a point transect survey with a NSEW neighbourhood, stepping outside of an animal cluster, and so not adding any additional observations to the survey. If the adaptive survey is configured incorrectly then adaptive efficiency may be less than a conventional survey. It appears that the adaptive approaches are often sensitive to minor alterations in the configuration or the population spatial characteristics so that a small change could potentially switch from a gain in efficiency to a loss. Certain configurations, such as the PB method with a zigzag pattern, are more robust, and this is probably linked to the ability to increase the effort in small increments rather than as an integer multiple of the sampling effort for a single point of transect leg.

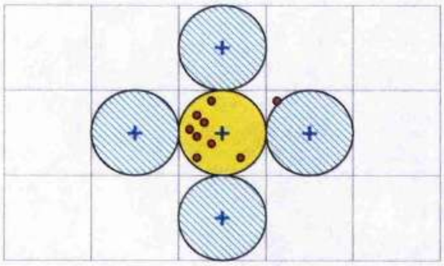


Figure 6.1: Adaptively sampled area, for a point transect survey, is much bigger than the actual spatial cluster of animals and so fails to increase observations by a worthwhile amount. The initial sampling plot is shown as a solid (yellow) circle and the adaptive sampling plots as (blue) cross-hatched circles.

With Thompson's adaptive methods, the total surveying effort is unknown at the start of the survey. For surveys where the adaptive component adds significant extra effort, this may be impractical, as many surveys are likely to be working with monetary and/or time constraints. In Chapter 3 we outlined approaches to constrain the total effort used, but these still required some kind of estimate as to the total adaptive effort that would be required and so were susceptible to marked increases in the actual surveying effort used. Brown and Manly (1998) have investigated an adaptive cluster sampling approach in which an upper limit on the number of units to sample is set prior to commencing the survey. Initial sample units are selected sequentially, and as each unit is selected and sampled any associated adaptive units are added. This continues until the cumulative total number of units sampled exceeds the preset sample size limit, or there are no more initial units to select. In this way the total number of units can only exceed the preset limit by a maximum of the number of adaptive units included by the last selected initial unit. However the sequential selection introduces some bias into the abundance estimate, although it is suggested that in some situations this can be estimated by bootstrapping. In addition, the method requires the initial units to be selected at random which could lead to convoluted travel between sampled units and an excess of off-effort travel.

The order in which transect legs are surveyed in a line transect survey must be considered. It is likely that there will be a trade off between following the most efficient survey path and ensuring a detected object has not moved out of the area covered by the adaptive legs. This is not restricted to line transects, and the comments also apply, although to a lesser extent, to point transects. Consider an adaptive line transect survey using a Parallel neighbourhood in a SIS design, with a short transect leg compared to the off-effort travel incurred moving to the parallel adaptive transect. Then if the frequency of adaptive triggers is high, it may use less off-effort travel, to complete the initial transect legs first before coming back down one side to perform all the adaptive sections on that side and then going back up the other side to complete all the adaptive legs there. Figure 6.2 shows this for a simplistic survey where the adaptive neighbourhood is restricted to a single transect leg on either side of the initial transect leg. Following this path will minimise the off-effort travel, but may not be the optimum route if there is movement of animals. If the animal movement is slow compared to the size of each transect leg sampling strip, this should not be a significant issue. However, if the speed is such that the cluster of animals could have moved out of an adaptive transect sampling strip before it has been surveyed, then the method may fail to add observations and so not improve efficiency. Conversely the animals that trigger the adaptive effort may have moved into the adaptive units. To reduce the possibility of either of these circumstances it is recommended, for fast moving animals, that adjacent adaptive transect legs are surveyed as soon as possible after its related trigger. Thus the size of the sampling units and their spacing is also a factor in how the survey performs, and it is preferable to keep the design such that off-effort travel is minimised.

The PB method effort factor calculation allocates effort as a proportion of the survey remaining. This means there is little capacity to accommodate lost effort towards the end of the survey. Thus in practice it may be preferable to add a constant to the nominal effort (say an additional 5-10%) to provide headroom to cope with any loss of surveying time in the final stages of the survey. Alternatively the survey could just switch to a conventional mode when the additional effort runs out, in which case it will overrun the survey effort/time limit.

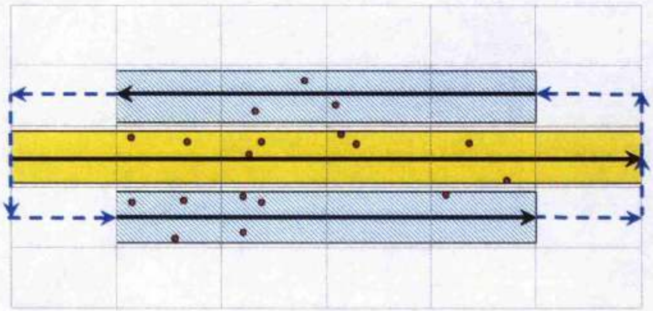


Figure 6.2: Following the most efficient survey path, may involve completing the initial transect legs first, then returning on one side to complete adaptive sections before crossing over and completing the adaptive transect legs on the other side. Off-effort is shown as a dashed line.

6.3 Future Developments

Areas warranting future development have been discussed at the end of each chapter. These focus on the need to identify suitable trigger conditions and adaptive patterns to optimise survey efficiency. Underlying this is the need to model the spatial distribution of the population being studied. This would be particularly useful for the PB method, as if a fixed total effort is used, then points or transects are not independent due to conditioning on the effort factors, and so cannot be re-sampled for a bootstrap. Modelling the population distribution would provide not only a mechanism for bootstrapping, but also enable enhanced simulation so that future surveys can be iteratively adjusted to improve their efficiency.

There has been limited simulation of the adaptive distance sampling based on Thompson's methods and a more detailed examination would assist in understanding where the approach will bring most benefit. Further evaluation of the PB method is also required, particularly with respect to the use of covariates in the detection function estimate and using effort factors less than one to account for poor coverage of an area.

Ideally a cost model should be produced so that the appropriate method, adaptive or conventional sampling, and suitable adaptive surveying parameters can be selected. Thompson considers cost models for his methods (Thompson, 1992: p275; Thompson, 1994; Thompson, 1996: p 149-161), but this is at a basic level and does not, for example, allow the impact of different trigger values to be compared.

6.4 Combining Conventional and Adaptive Surveys

When survey data have been gathered over a number of years, it may be desirable to continue to produce a conventional abundance estimate to provide a comparison over time. We close by considering how conventional and adaptive distance surveys can be combined such that both conventional and adaptive estimates are produced.

Extracting a conventional estimate from an adaptive design tends to be more straightforward for point transects than for line transects, and so we will deal with that first. For adaptive point transect surveys, using either the Thompson-based or the PB method, then it is purely a matter of dropping the adaptive points and analysing the remaining initial points with a conventional estimator. If comparison with previous survey data is required, it may be preferable to keep the number of initial points the same as previous surveys, and thus additional resource is going to be required for the adaptive surveys.

We now consider line transects using ideas originating from a request by a research unit interested in applying adaptive sampling to aerial surveys of small cetaceans such as narwhal (*Monodon monoceros*) and beluga (*Delphinapterus leucas*). As a result the discussion is focussed on this type of survey, although the approach is not limited to this alone. The research unit had experienced high levels of spatial clustering of animals and a large proportion of sightings were typically encountered on one or two transects. However the location of the aggregations was not sufficiently predictable to allow pre-stratification of the survey. The key requirements were for the survey to be practical for an aerial survey flown with GPS navigation and that any adaptive transect legs should be in addition to conventional (straight line transects), allowing conventional abundance estimates to be made and compared to previous survey results using the same initial transect design.

The outline survey method proposed was based on a SIS design, with primary units of transects made up of a number of contiguous transect legs. A trigger condition is defined based on the encounter rate rising above some value. When this value is exceeded, two adaptive transect legs are added parallel to and on either side of the current transect leg. The approach for adding the parallel transects is shown in Figure 6.3, and is as follows:

(1) The aircraft flies on an initial transect. Upon meeting the trigger condition the start position is recorded.
(2) The aircraft continues flying along the transect, until the trigger condition is no longer true and an end position is recorded.
(3) The aircraft then turns to one side and flies back, surveying at a pre-determined distance parallel to the initial transect until it gets to the start position.
(4) At this point the aircraft then turns across to the opposite side of the initial transect, and flies in the original direction, surveying at a pre-determined distance parallel to the initial transect.
(5) At the end position the aircraft then returns to the main transect and continues surveying.

On the next trigger of an adaptive section, the aircraft first turns to the opposite side to the previous adaptive section.

This fits easily with the PB method, as that has no requirement to divide the areas up into segments, and the parallel adaptive transects can be run between any two random positions on the initial trackline. However it is more difficult to apply the Thompson-based methods to this design, which typically sample fixed-size quadrats.

Overlaying the area with a randomly located grid of rectangles for the Thompson based methods provides a couple of advantages. The trigger condition is easier to measure, as it simply becomes a count between two specific locations or points in time, rather than say a rolling count in the last 15 minutes as was done for the experimental harbour porpoise survey. Dependent on the size of the rectangle, the cluster of animals is more likely to be located at some location within it, rather than being clipped at the edge of a start or stop position as may happen if using a rolling encounter rate to trigger. Thus there is less likelihood that the animals may have moved out of the area by the time the observer returns to sample the adaptive transect legs.

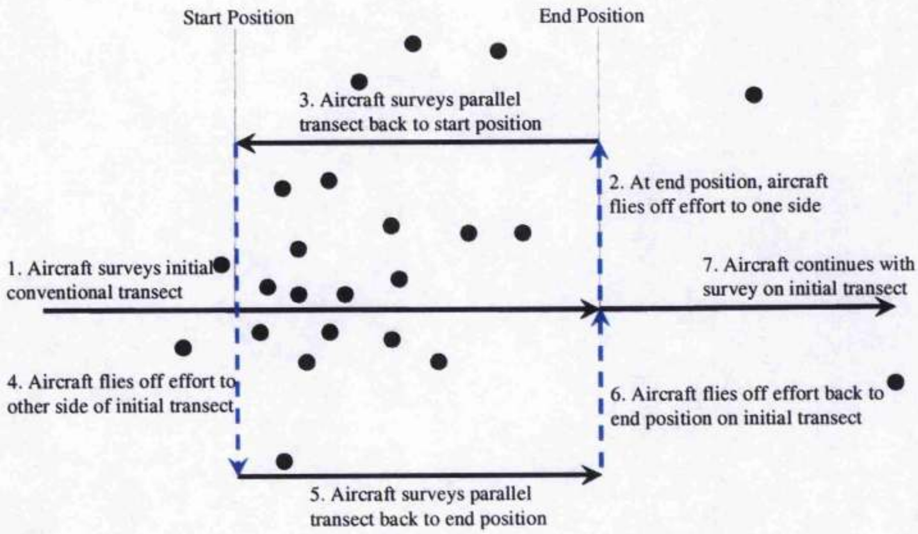


Figure 6.3: Basic survey procedure. Aircraft surveys the initial transect until the trigger condition is met, and records the start position. Aircraft continues surveying transect until trigger condition ceases to be true, at which point the end position is recorded

If the survey region is overlaid with a randomly located grid of units, it is possible to produce conventional, Thompson-based and PB method estimators of the population abundance. For the conventional estimate you simply drop all adaptive transect legs and only analyse the initial transect legs. For the Thompson-based methods, the estimators of Chapter 3 are used; and for the PB method, the estimators of Chapter 4 used.

As a large proportion of sightings are limited to one or two transects in these surveys, it is suggested that adaptive legs are allowed to continue adapting. The intent is for around 10% of the survey effort to be adaptive. Thus by reviewing previous survey data and additional strip transect photographs, it should be possible to estimate a suitable trigger condition. If this proves too low and there is excessive adaptive effort being used, then the trigger can be revised to a higher value, mid-survey. In this case any sampled adaptive transect legs that no longer meet the new trigger condition are dropped from the analysis (although it may still be possible to use the sightings in the detection function estimation).

For the PB method, it will not be possible to make the effort factor a function of the available effort remaining. For each initial transect leg the effort factor will be an integer multiple, simply calculated as one plus the total number of adaptive transects legs associated with the triggering initial transect leg.

If the analysis using the Thompson-based methods is not required then it would be possible to make the effort factor a function of the remaining available effort. However the existing effort function would work slightly differently. If the adaptive trigger adds a parallel transect leg on each side of the initial transect leg, then for the first trigger on an initial transect leg, the effort factor is 3; two adaptive legs plus one initial leg. After this for each adaptive leg that meets the trigger condition only one further transect leg is added, as there will already be a transect leg on the other side, and so the effort factor is incremented by one each time. Thus the effort factor function needs to be modified so it returns a Boolean value of true or false. If false is returned, then no further adaptive legs are added. A result of true signifies that two adaptive legs are added if triggered by an initial transect, or a single adaptive leg added if the trigger is from an adaptive leg. As the addition of parallel transects also adds off-effort travel it may be considered worthwhile to include this in the calculation, perhaps as a proportion of surveying effort if it is possible to traverse off-effort sections at a different speed.

Appendix A

RATS

A.1 Introduction

A simulation computer program RATS (Restricted Adaptive Transect Sampling) was developed to compare adaptive sampling with fixed effort and conventional line transect methods. It has a graphical user interface, to allow the visual representation of populations and surveys, and was developed in C++ to run on Personal Computers running the Microsoft Windows operating system. Summarised results are displayed on the screen, whilst detailed output and analysis is produced either as a plain text report or as rows of data (for multiple simulations) to a comma separated variable (CSV) file so that it can be loaded into a spreadsheet or database for further analysis.

The program can: create a simulated population, based on a Poisson cluster process (Diggle, 1983); simulate a conventional or adaptive (with fixed effort) line transect survey; analyse survey results; and where required automatically run and analyse a large number of simulations. The program provided a useful basis for other distance sampling simulations and so was extended to simulate: conventional point transect surveys; and adaptive point and line transect surveys based on Thompson's methods. However the program has less functionality in these cases and the results analysis is not fully automated. An example screen with summary results for a adaptive line transect survey with fixed effort is shown in Figure A.1

RATS simulates a population within a rectangular frame, where the height and width of the frame are configuration options, but are typically set to 100 by 100 units. Conventional and/or adaptive line and point transect surveys can then be run on the simulated population. For line transects, the nominal transects run from left to right across the population frame. The results from a simulated survey are analysed using the formulae developed in this thesis. Some sample simulations are shown in Appendix B.

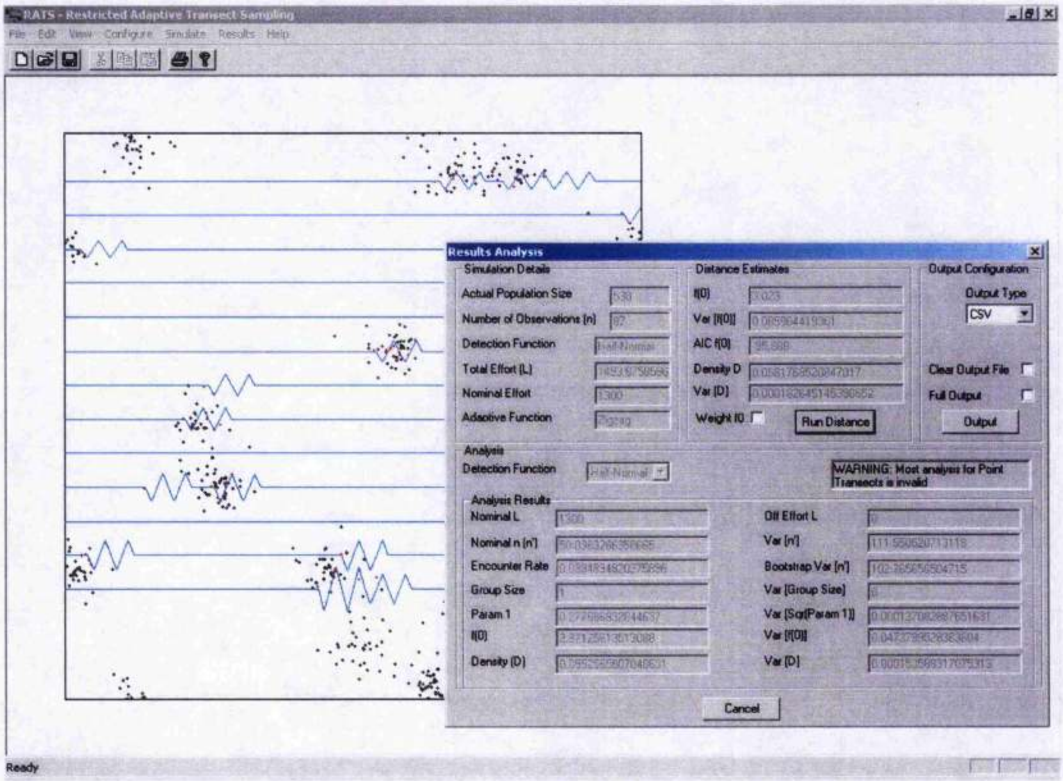


Figure A.1: Example screen from RATS, showing summary analysis report in the foreground and a simulated adaptive line transect survey in the background.

As well as simulating surveys, the program also has the capability to read in survey results, and analyse them using the same analysis engine as the simulations.

The program uses a unit coordinate system and thus all results are relative to these units. The user can then translate these to their measurement unit of choice, such as kilometres or Nautical miles.

A.1.1 Population Simulation

Populations are simulated using randomly located parent clusters, each of which consists of a number of objects (animal groups) distributed around the parent cluster centre. Each group has a fixed size of 1, that is, the simulated populations comprise individual animals, each of which belongs to a loose spatial cluster of animals. Thus the program does not exercise the group size estimators fully, although the formulae are implemented in the analysis part of the program The default set-up provides a

Poisson cluster process (Diggle 1983: p55). The set-up screen for the population parameters is shown in Figure A.2.

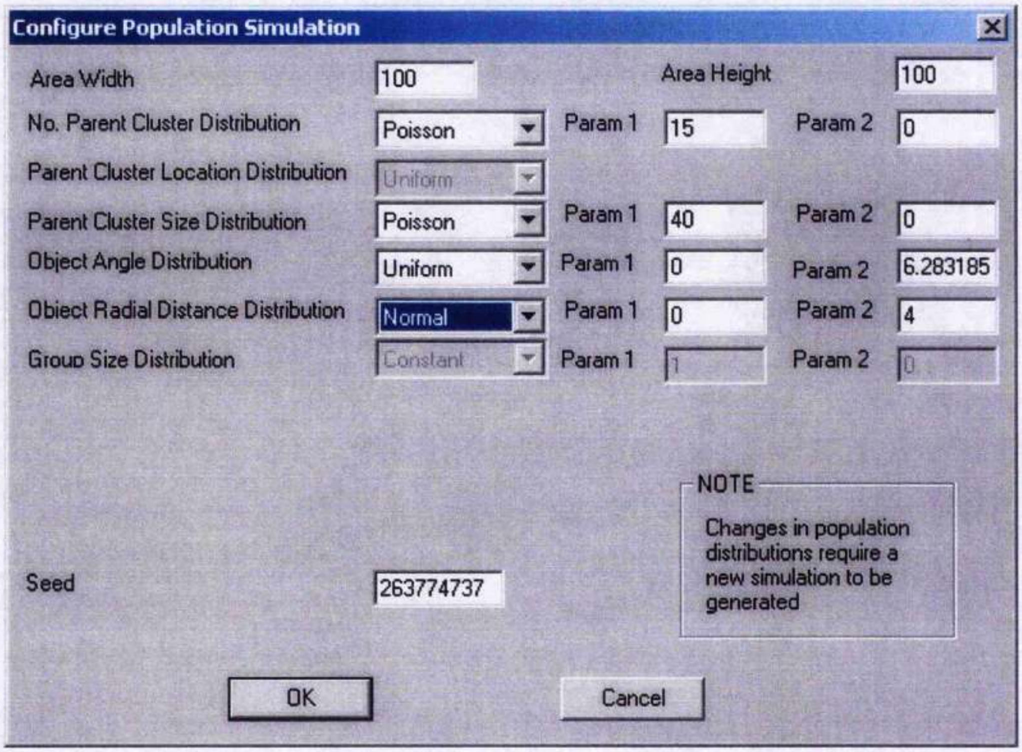


Figure A.2: Population configuration screen.

The population is created as follows:

(1) The number of parent clusters is simulated using an appropriate distribution selected by the user from the options available.

(2) For each parent cluster the following is then performed:

(i) The number of animals within the parent cluster is simulated, using a distribution selected by the user from the list available.

(ii) The centre of the parent cluster is simulated using continuous uniform variates between 0 and area height or width (default 100 units) to simulate its horizontal and vertical position within the main population frame. Thus there is no gradient in the simulated population densities in these simulations.

(iii) For each animal in a parent cluster, its position, relative to the parent cluster centre, is simulated using a radial angle and distance. The angle is simulated using a continuous uniform variate between 0 and 2π. The radial distance from the centre to the animal is simulated using a distribution selected by the user. This relative position is then converted to the animal's actual position relative to the population frame. If this position lies within the frame then the animal is included in the population. If the animal lies outside the frame, then the distance to the animal is 'wrapped around' to the opposite edge. This is performed both horizontally and vertically as necessary, and is repeated until the animal lies inside the population frame.

The components of a simulated population are summarised in Figure A.3. The transect start positions are simulated to generate randomly positioned lines parallel to the top and bottom edges of the population frame. The total number of transects can be set by the user.

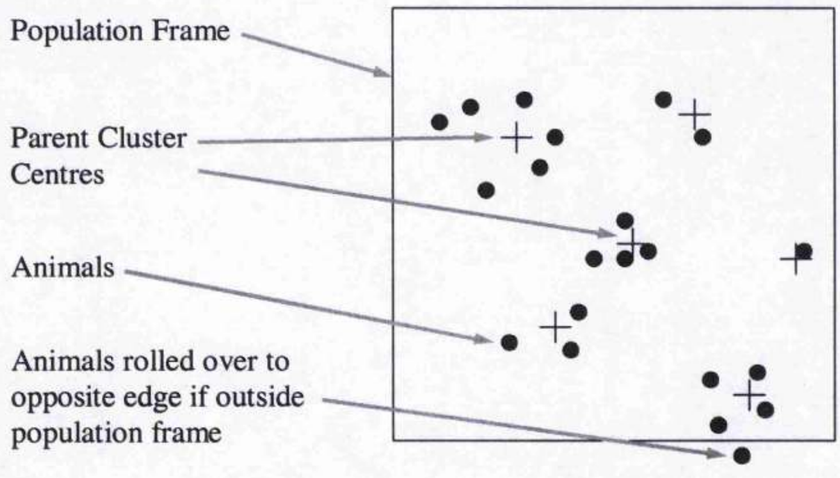


Figure A.3: Populations are simulated in clusters within a population frame. Animals lying outside the frame are repeatedly 'wrapped around' until they lie within the frame.

A.1.2 Survey Simulation

The detection function is simulated using a half-normal distribution, with no adjustment terms, scaled so that the probability of detection on the line is 1. Thus, the detection function is

$$g(x) = e^{-x^2/2\sigma^2}$$

The parameter σ is a configurable through the setup options. For line transects a different value for σ can optionally be set on adaptive sections, typically use to simulate increased observer awareness following a sighting. A third setting of the parameter σ can also be optionally set, to be used for configurable amount of effort, and this is typically used to simulate a reduced effective strip half-width, due to example bad weather.

The truncation half-width is also configurable. As is the trigger value, which always relates to the encounter rate for either the previous step length in the case of a line transect survey, or previous point for a point transect survey. An example of the configuration screen for a line transect survey is shown in Figure A.4.

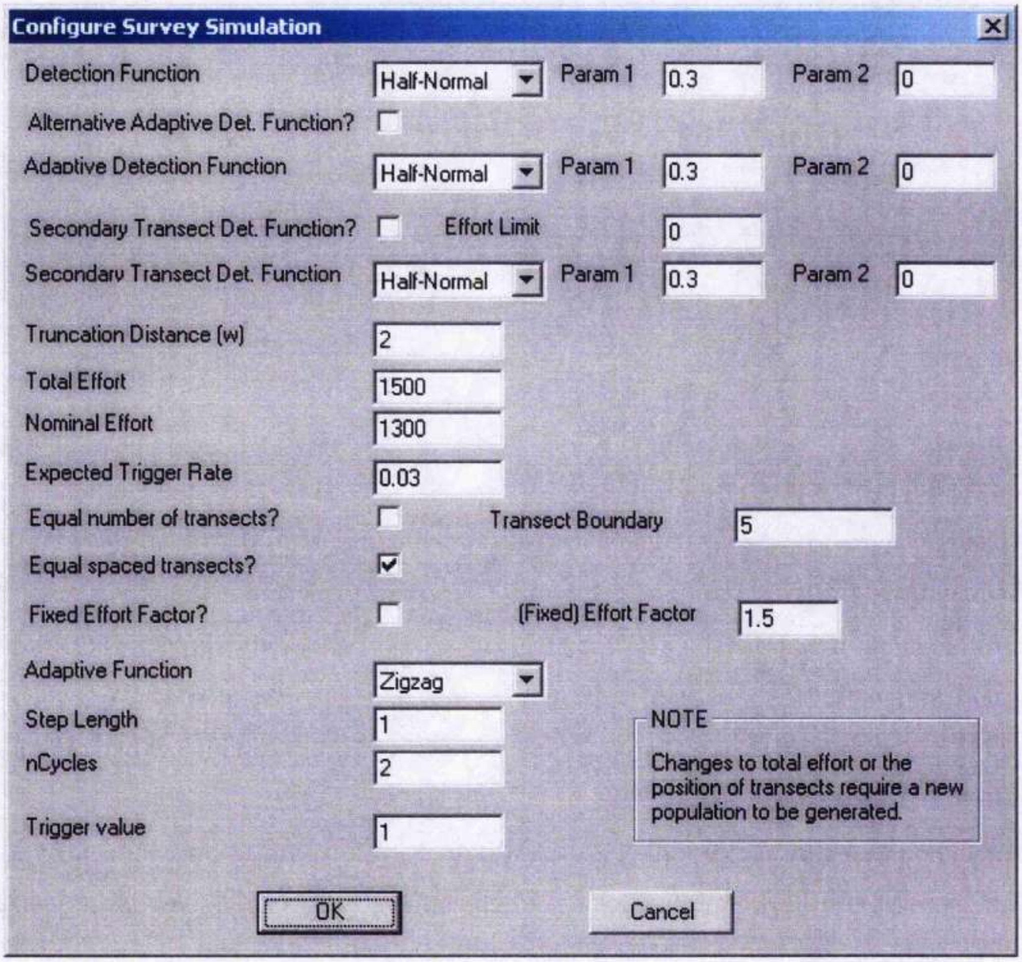


Figure A.4: Line transect survey simulation configuration screen.

Line Transect Survey Simulation

Transects are run from left to right, horizontally across the population frame. The transect start positions are can be either randomly or systematically selected on the vertical axis. A horizontal buffer zone is configurable, across both the top and the bottom of the population frame, to prevent reduce edge effects where transects may adapt over the edge of the population frame. Each transect is traversed in horizontal steps, with the step size configured by the user. For a conventional line transect survey, at each step an area is sampled using a rectangle centred on the transect. With an adaptive survey, using a zigzag pattern, the rectangle becomes a parallelogram for the zigzag sections. The perpendicular offset to the edge of the rectangle or parallelogram, on each side of the transect, is the value *w* (truncation half-width) specified by the user (Figure A.5).

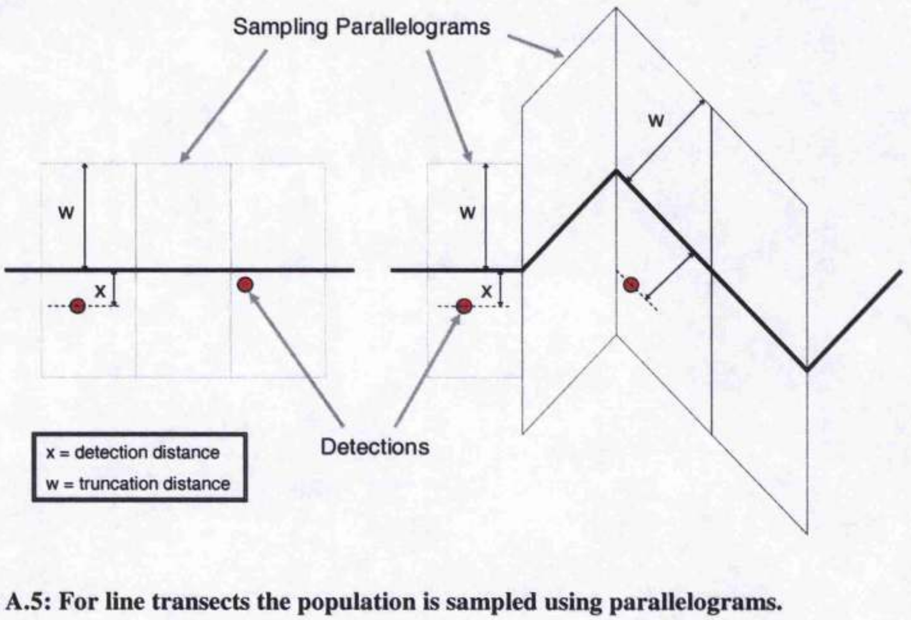


Figure A.5: For line transects the population is sampled using parallelograms.

For any animal within the parallelogram, detection is simulated using a half-normal detection function as already described.

A number of adaptive patterns are available, including a zigzag pattern, a hounds-tooth pattern (as described in Chapter 5) and simply running parallel transects. In each case the number of times the pattern is repeated can be configured. With any of the patterns, a trigger on the last step length of the pattern will trigger the adapting to continue. Each adaptive transect starts in conventional survey mode, even if the survey was adapting at the end of the previous transect.

Automated runs are able to simulate a number of populations in sequence and, for each population, an adaptive and a conventional line transect survey can be run. For both surveys, the same transect start points are used but the detection process is simulated independently.

The nominal length of each transect can also be specified. There are two options available:

(1) The conventional and adaptive surveys use transects of equivalent nominal length. In this case, the transects run the full width of the survey area for both survey types. This means there are fewer transects in the adaptive survey than the conventional survey.

(2) The adaptive survey uses a nominal length which is scaled such that there is an equivalent number of transects for the two types of survey. When the adaptive transect length is scaled down, the transects all start from the left hand side of the area. In our simulations, this is not biased as there is no gradient in the simulated populations. In a real survey, such an option would not be used, but it was implemented here to aid comparison between adaptive and conventional sampling.

Point Transect Survey Simulation

Point transects locations are simulated using a systematic grid, with equal inter-point spacing both horizontally and vertically. The number of points used is changed by modifying the program, although the default is 100. If a boundary width is set, to reduce edge effects, then in this case the boundary zone is along all four edges of the population frame (top, bottom, left and right).

Points are sampled using the simulated half-normal detection function, although unlike line transects, there are not options to use alternative values for σ to simulate increased observer awareness and bad weather. The adaptive pattern is with a NSEW neighbourhood, as defined in Chapter 3.

At the time of writing very limited point transect analysis is performed by the program, and for most cases the survey results are output to a text file for the user to calculate estimates using another tool. For the simulation in this paper, the point transect estimates were produced using the Microsoft Excel spreadsheet and a number of custom written macros.

A.1.3 Analysis

RATS can use a built in mechanism to estimate the detection function, in which case there is no model selection as the observation data is fitted to a half-normal function assuming no truncation of the data. i.e. For a line transect survey

$$\hat{f}(0) = \sqrt{2/\pi\hat{\sigma}^2}$$

where

$$\hat{\sigma}^2 = \sum_{i=1}^{n} x_i^2 / n$$

and the x_i are the recorded detection distances for the n observations. The variance is estimated by

$$\hat{V}\{\hat{f}(0)\} = \frac{\{f(0)\}^2}{2n}$$

Alternatively DISTANCE 2.2 (Laake *et al.*, 1994) is used to fit the detection function, and the choice of models to select from is configured using a DISTANCE input file. Using DISTANCE is preferable as it is a more representative of a real survey where the true detection is unknown. If the inbuilt detection function is used, then the user needs to be conscious of choosing a suitable truncation distance. If the truncation half-width is set too small for the value of σ used in the simulations, then there may be large bias in the estimate.

The probability of missing a detection outside the truncation width can be calculated as follows:

The cumulative distribution function (cdf) for the half-normal detection function is given by

$$G(x) = \int_0^x e^{-x^2/2\sigma^2}\, dx$$

The cdf for a normal distribution, with a mean μ and standard deviation σ is given by

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-(x-\mu)^2/2\sigma^2}\, dx$$

as $\lim_{x \to \infty} F(x) = 1$, and the function is symmetric, then

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_0^{x} e^{-(x-\mu)^2/2\sigma^2}\, dx + 0.5$$

so that, setting the mean equal to zero ($\mu = 0$), then

$$\begin{aligned} G(x) &= \sqrt{2\pi}\sigma \cdot \{F(x) - 0.5\} \\ &= \sqrt{2\pi}\sigma \cdot \{\Phi(x/\sigma) - 0.5\} \end{aligned}$$

where

$\Phi(.)$ is the cdf for the standard normal function.

Thus probability that an object is detected within a truncation half-width, w, is

$$P(W \le w) = \frac{\int_0^{w} e^{-w^2/2\sigma^2}\, dw}{\int_0^{\infty} e^{-w^2/2\sigma^2}\, dw} = \frac{\sqrt{2\pi}\sigma \cdot \left(\Phi\left(\frac{w}{\sigma}\right) - \frac{1}{2}\right)}{\sqrt{\frac{\pi\sigma^2}{2}}} = 2 \cdot \left(\Phi\left(\frac{w}{\sigma}\right) - \frac{1}{2}\right)$$

Thus the probability that an object would have been detected beyond the truncation half-width w, is

$$P(W > w) = 1 - P(W \le w) = 1 - 2 \cdot \left(\Phi\left(\frac{w}{\sigma}\right) - \frac{1}{2}\right)$$

This can be expressed in Microsoft Excel as

=1-2*(NORMSDIST(w/σ)-0.5)

where, w and σ are entered as the appropriate truncation half-width and standard deviation values being simulated.

Appendix B

Simulation Results

This appendix contains the results for the adaptive line transect with fixed effort simulations in Chapter 4. In addition example conventional and adaptive simulations are shown for each of the 3 basic population types used.

The results are summaries of the 1000 simulation comparisons for the five types of simulation undertaken:

1) CSR population;
2) Clustered population;
3) Highly Clustered population;
4) Highly Clustered population with alternative $f(0)$ on adaptive sections;
5) Highly Clustered population with 400 units of bad weather.

Each summary contains 6 columns. The first column is the name or description of the value. The second and third columns are the mean and, where appropriate, the standard deviation for the adaptive value over the 1000 simulations. If the value is a range then the second and third columns are instead the low and high values respectively. The context can be determined from the name. The fourth and fifth columns have the same representation as the second and third, but contain the conventional simulation values. Finally for selected values the sixth column contains an improvement percentage, where improvement is measured as the conventional value divided by the adaptive value.

The table has 4 main sections. There is a GENERAL section which contain information about the survey and the population, such as the population size. Then there are sections for ENCOUNTER RATE, f(0) and DENSITY, which contain a summary of the information for the three estimators.

The majority of the names used should be self explanatory. The detection function was fitted both using the inbuilt mechanism in RATS and using DISTANCE, although results were only used from the DISTANCE estimation as it included variance from model selection. Where fitted by DISTANCE then the name is prefixed by 'DIST', for example 'DIST est.f(0)' and where fitted by RATS they are prefixed by 'RATS; This is also true of any resulting estimators, such as 'DIST est.D' and 'Bias RATS est.D'. For all models a weighted maximum equations, 4.22 and 4.23 were also used to estimate $f(0)$ and these are signified by 'WL_', for example 'RATS est.WL_D'. Percent relative bias is referred to as PRB, the Root Mean Square Errors as RMSE and Confidence Interval as CI.

B.1 CSR

	ADAPTIVE		CONVENTIONAL		
	Mean (or Low)	SD (or High)	Mean (or Low)	SD (or High)	Improvement
GENERAL					
Total Effort	1498.6	1.63	1500.0	0.00	
Nominal Effort	1300	0	1300	0	
Population Size	600.0	0.00	600.0	0.00	
Pop. Index of Dispersion	1.03	0.353	1.03	0.353	
No. Observations	67.47	7.621	67.78	7.783	
No. Nom. Observations	58.54	6.762	67.78	7.783	
ENCOUNTER RATE					
True Encounter Rate	0.0451	6.832E-09	0.0451	6.832E-09	
Est.E[e|L']	0.0450	0.005201	0.0452	0.005189	
Bias est.E[e|L']	-0.0000877	0.005201	0.0000680	0.005189	
RMSE	0.00520		0.00519		0.998
Est.V[est.E[e|L']]	0.0000313	1.325E-05	0.0000300	1.155E-05	0.959
n below/above 95% CI	30	32	30	27	
PRB	-0.19%		0.15%		
PRB -/+	-0.91%	0.52%	-0.56%	0.86%	
f(0)					
True Overall f(0)	2.660		2.660		
RATS est.f(0)	2.684	0.2340	2.672	0.2278	
Bias RATS est.f(0)	0.0248	0.2340	0.0122	0.2278	
RMSE	0.235		0.228		0.969
RATS est.V[est.f(0)]	0.0546	0.01218	0.0538	0.01184	0.986
n below/above 95% CI	29	13	25	16	
PRB	0.93%		0.46%		
PRB -/+	0.39%	1.48%	-0.07%	0.99%	
DIST est.f(0)	2.641	0.3297	2.643	0.3357	
Bias DIST est.f(0)	-0.0182	0.3297	-0.0166	0.3357	
RMSE	0.330		0.336		1.018
DIST est.V[est.f(0)]	0.0728	0.06619	0.0751	0.06911	1.032
n below/above 95% CI	29	132	41	132	
PRB	-0.69%		-0.63%		
PRB -/+	-1.45%	0.08%	-1.41%	0.16%	
RATS est.WL_f0	2.688	0.2426	0.891	0.2278	
Bias RATS est.WL_f0	0.0279	0.2426	0.0122	0.2278	
RMSE	0.244		0.228		0.934
RATS est.V[est.WL_f0]	0.0631	0.01460	0.0538	0.01184	0.852
n below/above 95% CI	23	11	25	16	
PRB	1.05%		0.46%		
PRB -/+	0.48%	1.61%	-0.07%	0.99%	
DENSITY					
True Density	0.06	0	0.06	0	
RATS est.D	0.0604	0.008487	0.0603	0.008459	
Bias RATS est.D	0.000410	0.008487	0.000345	0.008459	
RMSE	0.00849		0.00846		0.996
RATS est.V[est.D]	0.0000843	0.00003024	0.0000809	0.00002532	0.960
n below/above 95% CI	23	23	22	21	
PRB	0.68%		0.58%		
PRB -/+	-0.194%	1.560%	-0.299%	1.449%	
DIST est.D	0.0594	0.009802	0.0597	0.010133	
Bias DIST est.D	-0.000567	0.009802	-0.000294	0.010133	
RMSE	0.00981		0.01013		1.032
DIST est.V[est.D]	0.0000919	0.00004978	0.0000910	0.00004899	0.990
n below/above 95% CI	28	57	34	49	
PRB	-0.95%		-0.49%		
PRB -/+	-1.958%	0.067%	-1.536%	0.557%	
RATS est.WL_D	0.06048	0.008628	0.06035	0.008459	
Bias RATS est.WL_D	0.000481	0.008628	0.000345	0.008459	
RMSE	0.00864		0.00846		0.980
RATS est.V[est.WL_D]	8.87E-05	3.140E-05	8.09E-05	2.532E-05	0.912
n below/above 95% CI	21	19	22	21	
PRB	0.80%		0.58%		
PRB -/+	-0.09%	1.69%	-0.30%	1.45%	

B.2 Clustered

	ADAPTIVE		CONVENTIONAL		
	Mean (or Low)	SD (or High)	Mean (or Low)	SD (or High)	Improvement
GENERAL					
Total Effort	1492.8	8.52	1500.0	0.00	
Nominal Effort	1300	0	1300	0	
Population Size	601.6	94.06	601.6	94.06	
Pop. Index of Dispersion	12.29	4.385	12.29	4.385	
No. Observations	79.27	15.975	67.55	13.828	
No. Nom. Observations	58.48	12.609	67.55	13.828	
ENCOUNTER RATE					
True Encounter Rate	0.0452	7.073E-03	0.0452	7.073E-03	
Est.E[e|L']	0.0450	0.009699	0.0450	0.009219	
Bias est.E[e|L']	-0.0002573	0.006482	-0.0002132	0.006178	
RMSE	0.00648		0.00618		0.953
Est.V[est.E[e|L']]	0.0000709	3.589E-05	0.0000704	3.425E-05	0.993
n below/above 95% CI	12	17	4	12	
PRB	-0.57%		-0.47%		
PRB -/+	-1.46%	0.32%	-1.32%	0.38%	
f(0)					
True Overall f(0)	2.660		2.660		
RATS est.f(0)	2.694	0.2367	2.693	0.2388	
Bias RATS est.f(0)	0.0341	0.2367	0.0332	0.2388	
RMSE	0.239		0.241		1.008
RATS est.V[est.f(0)]	0.0482	0.01423	0.0566	0.01675	1.175
n below/above 95% CI	49	22	33	17	
PRB	1.28%		1.25%		
PRB -/+	0.73%	1.83%	0.69%	1.80%	
DIST est.f(0)	2.630	0.3195	2.645	0.3408	
Bias DIST est.f(0)	-0.0292	0.3195	-0.0145	0.3408	
RMSE	0.321		0.341		1.063
DIST est.V[est.f(0)]	0.0587	0.05385	0.0743	0.07218	1.265
n below/above 95% CI	53	162	38	129	
PRB	-1.10%		-0.55%		
PRB -/+	-1.84%	-0.35%	-1.34%	0.25%	
RATS est.WL_f0	2.691	0.2470	0.898	0.2388	
Bias RATS est.WL_f0	0.0313	0.2470	0.0332	0.2388	
RMSE	0.249		0.241		0.968
RATS est.V[est.WL_f0]	0.0656	0.02054	0.0566	0.01675	0.863
n below/above 95% CI	22	13	33	17	
PRB	1.18%		1.25%		
PRB -/+	0.60%	1.75%	0.69%	1.80%	
DENSITY					
True Density	0.0601647	0.009405563	0.0601647	0.009405563	
RATS est.D	0.0606	0.013937	0.0606	0.013273	
Bias RATS est.D	0.000386	0.010213	0.000413	0.009763	
RMSE	0.01022		0.00977		0.956
RATS est.V[est.D]	0.0001531	0.00007517	0.0001555	0.00007006	1.016
n below/above 95% CI	13	12	12	10	
PRB	0.64%		0.69%		
PRB -/+	-0.411%	1.693%	-0.320%	1.692%	
DIST est.D	0.0591	0.014429	0.0596	0.014397	
Bias DIST est.D	-0.001034	0.011100	-0.000602	0.011179	
RMSE	0.01114		0.01119		1.004
DIST est.V[est.D]	0.0001529	0.00008345	0.0001606	0.00008940	1.050
n below/above 95% CI	11	35	8	23	
PRB	-1.72%		-1.00%		
PRB -/+	-2.862%	-0.575%	-2.153%	0.150%	
RATS est.WL_D	0.06049	0.013997	0.06058	0.013273	
Bias RATS est.WL_D	0.000325	0.010318	0.000413	0.009763	
RMSE	0.01032		0.00977		0.947
RATS est.V[est.WL_D]	1.61E-04	7.614E-05	1.56E-04	7.006E-05	0.966
n below/above 95% CI	12	9	12	10	
PRB	0.54%		0.69%		
PRB -/+	-0.52%	1.60%	-0.32%	1.69%	

B.3 Highly Clustered

	ADAPTIVE		CONVENTIONAL		
	Mean (or Low)	SD (or High)	Mean (or Low)	SD (or High)	Improvement
GENERAL					
Total Effort	1476.4	24.27	1500.0	0.00	
Nominal Effort	1300	0	1300	0	
Population Size	605.4	156.62	605.4	156.62	
Pop. Index of Dispersion	31.22	11.326	31.22	11.326	
No. Observations	95.60	25.821	68.65	20.888	
No. Nom. Observations	58.65	18.639	68.65	20.888	
ENCOUNTER RATE					
True Encounter Rate	0.0455	1.178E-02	0.0455	1.178E-02	
Est.E[e|L']	0.0451	0.014338	0.0458	0.013925	
Bias est.E[e|L']	-0.0004125	0.008035	0.0002372	0.007519	
RMSE	0.00804		0.00752		0.935
Est.V[est.E[e|L']]	0.0001355	8.329E-05	0.0001402	8.279E-05	1.035
n below/above 95% CI	4	11	3	6	
PRB	-0.91%		0.52%		
PRB -/+	-2.00%	0.19%	-0.50%	1.54%	
f(0)					
True Overall f(0)	2.660		2.660		
RATS est.f(0)	2.706	0.2065	2.690	0.2543	
Bias RATS est.f(0)	0.0468	0.2065	0.0302	0.2543	
RMSE	0.212		0.256		1.209
RATS est.V[est.f(0)]	0.0421	0.01676	0.0596	0.02888	1.417
n below/above 95% CI	44	10	43	20	
PRB	1.76%		1.14%		
PRB -/+	1.28%	2.24%	0.54%	1.73%	
DIST est.f(0)	2.661	0.3038	2.629	0.3612	
Bias DIST est.f(0)	0.0010	0.3038	-0.0310	0.3612	
RMSE	0.304		0.362		1.193
DIST est.V[est.f(0)]	0.0547	0.05581	0.0738	0.08002	1.349
n below/above 95% CI	63	131	50	140	
PRB	0.04%		-1.17%		
PRB -/+	-0.67%	0.74%	-2.01%	-0.32%	
RATS est.WL_f0	2.704	0.2270	0.897	0.2543	
Bias RATS est.WL_f0	0.0448	0.2270	0.0302	0.2543	
RMSE	0.231		0.256		1.106
RATS est.V[est.WL_f0]	0.0709	0.03315	0.0596	0.02888	0.841
n below/above 95% CI	19	6	43	20	
PRB	1.68%		1.14%		
PRB -/+	1.15%	2.21%	0.54%	1.73%	
DENSITY					
True Density	0.0605437	0.015661554	0.0605437	0.015661554	
RATS est.D	0.0611	0.020017	0.0615	0.019466	
Bias RATS est.D	0.000529	0.011954	0.000952	0.011859	
RMSE	0.01196		0.01189		0.994
RATS est.V[est.D]	0.0002695	0.00016623	0.0002837	0.00016692	1.053
n below/above 95% CI	9	10	4	7	
PRB	0.87%		1.57%		
PRB -/+	-0.351%	2.097%	0.358%	2.786%	
DIST est.D	0.0600	0.020221	0.0602	0.020031	
Bias DIST est.D	-0.000522	0.012680	-0.000339	0.013097	
RMSE	0.01268		0.01309		1.032
DIST est.V[est.D]	0.0002677	0.00016938	0.0002836	0.00018513	1.059
n below/above 95% CI	6	21	4	25	
PRB	-0.86%		-0.56%		
PRB -/+	-2.161%	0.436%	-1.900%	0.782%	
RATS est.WL_D	0.06100	0.020024	0.06150	0.019466	
Bias RATS est.WL_D	0.000456	0.012044	0.000952	0.011859	
RMSE	0.01205		0.01189		0.987
RATS est.V[est.WL_D]	2.82E-04	1.697E-04	2.84E-04	1.669E-04	1.008
n below/above 95% CI	8	11	4	7	
PRB	0.75%		1.57%		
PRB -/+	-0.48%	1.99%	0.36%	2.79%	

B.4 Highly Clustered with Alternative Adaptive *f*(0)

	ADAPTIVE		CONVENTIONAL		
	Mean (or Low)	SD (or High)	Mean (or Low)	SD (or High)	Improvement
GENERAL					
Total Effort	1476.4	24.27	1500.0	0.00	
Nominal Effort	1300	0	1300	0	
Population Size	605.4	156.62	605.4	156.62	
Pop. Index of Dispersion	31.22	11.326	31.22	11.326	
No. Observations	114.23	30.501	68.65	20.888	
No. Nom. Observations	67.78	21.569	68.65	20.888	
ENCOUNTER RATE					
True Encounter Rate	0.0455	1.178E-02	0.0455	1.178E-02	
Est.E[e|L']	0.0521	0.016591	0.0458	0.013925	
Bias est.E[e|L']	0.0066123	0.009441	0.0002372	0.007519	
RMSE	0.01152		0.00752		0.653
Est.V[est.E[e|L']]	0.0001848	1.134E-04	0.0001402	8.279E-05	0.759
n below/above 95% CI	35	2	3	6	
PRB	14.52%		0.52%		
PRB -/+	13.24%	15.81%	-0.50%	1.54%	
f(0)					
True Overall f(0)	2.660		2.660		
RATS est.f(0)	2.210	0.1670	2.690	0.2543	
Bias RATS est.f(0)	-0.4501	0.1670	0.0302	0.2543	
RMSE	0.480		0.256		0.533
RATS est.V[est.f(0)]	0.0234	0.00963	0.0596	0.02888	2.542
n below/above 95% CI	0	764	43	20	
PRB	-16.92%		1.14%		
PRB -/+	-17.31%	-16.53%	0.54%	1.73%	
DIST est.f(0)	2.210	0.2372	2.629	0.3612	
Bias DIST est.f(0)	-0.4495	0.2372	-0.0310	0.3612	
RMSE	0.508		0.362		0.713
DIST est.V[est.f(0)]	0.0334	0.03302	0.0738	0.08002	2.209
n below/above 95% CI	1	676	50	140	
PRB	-16.90%		-1.17%		
PRB -/+	-17.46%	-16.35%	-2.01%	-0.32%	
RATS est.WL_f0	2.273	0.1697	0.897	0.2543	
Bias RATS est.WL_f0	-0.3867	0.1697	0.0302	0.2543	
RMSE	0.422		0.256		0.606
RATS est.V[est.WL_f0]	0.0434	0.02053	0.0596	0.02888	1.374
n below/above 95% CI	0	447	43	20	
PRB	-14.54%		1.14%		
PRB -/+	-14.93%	-14.14%	0.54%	1.73%	
DENSITY					
True Density	0.0605437	0.015661554	0.0605437	0.015661554	
RATS est.D	0.0576	0.018976	0.0615	0.019466	
Bias RATS est.D	-0.002904	0.011275	0.000952	0.011859	
RMSE	0.01164		0.01189		1.022
RATS est.V[est.D]	0.0002419	0.00015060	0.0002837	0.00016692	1.173
n below/above 95% CI	2	25	4	7	
PRB	-4.80%		1.57%		
PRB -/+	-5.952%	-3.643%	0.358%	2.786%	
DIST est.D	0.0577	0.019642	0.0602	0.020031	
Bias DIST est.D	-0.002844	0.012057	-0.000339	0.013097	
RMSE	0.01238		0.01309		1.058
DIST est.V[est.D]	0.0002490	0.00016019	0.0002836	0.00018513	1.139
n below/above 95% CI	2	27	4	25	
PRB	-4.70%		-0.56%		
PRB -/+	-5.931%	-3.463%	-1.900%	0.782%	
RATS est.WL_D	0.05922	0.019268	0.06150	0.019466	
Bias RATS est.WL_D	-0.001319	0.011418	0.000952	0.011859	
RMSE	0.01149		0.01189		1.035
RATS est.V[est.WL_D]	2.66E-04	1.591E-04	2.84E-04	1.669E-04	1.068
n below/above 95% CI	2	19	4	7	
PRB	-2.18%		1.57%		
PRB -/+	-3.35%	-1.01%	0.36%	2.79%	

B.5 Highly Clustered with 400 units of Bad Weather

	ADAPTIVE		CONVENTIONAL		
	Mean (or Low)	SD (or High)	Mean (or Low)	SD (or High)	Improvement
GENERAL					
Total Effort	1477.3	23.92	1500.0	0.00	
Nominal Effort	1300	0	1300	0	
Population Size	605.9	157.81	605.9	157.81	
Pop. Index of Dispersion	31.07	11.360	31.07	11.360	
No. Observations	84.62	24.471	59.78	19.325	
No. Nom. Observations	50.49	17.058	59.78	19.325	
ENCOUNTER RATE					
True Encounter Rate	0.0359	9.338E-03	0.0360	9.369E-03	
Est.E[e|L']	0.0388	0.013122	0.0399	0.012884	
Bias est.E[e|L']	0.0029825	0.007885	0.0038826	0.007570	
RMSE	0.00843		0.00850		1.009
Est.V[est.E[e|L']]	0.0001144	7.733E-05	0.0001213	7.494E-05	1.061
n below/above 95% CI	19	12	16	1	
PRB	8.32%		10.79%		
PRB -/+	6.96%	9.68%	9.49%	12.10%	
f(0)					
True Overall f(0)	3.380		3.369		
RATS est.f(0)	2.868	0.2982	2.899	0.3511	
Bias RATS est.f(0)	-0.5122	0.2982	-0.4696	0.3511	
RMSE	0.593		0.586		0.989
RATS est.V[est.f(0)]	0.0560	0.03737	0.0826	0.06100	1.475
n below/above 95% CI	2	597	3	439	
PRB	-15.16%		-13.94%		
PRB -/+	-15.70%	-14.61%	-14.58%	-13.29%	
DIST est.f(0)	2.925	0.4649	2.962	0.5505	
Bias DIST est.f(0)	-0.4549	0.4649	-0.4067	0.5505	
RMSE	0.650		0.684		1.052
DIST est.V[est.f(0)]	0.1114	0.12946	0.1742	0.18920	1.563
n below/above 95% CI	12	489	10	384	
PRB	-13.46%		-12.07%		
PRB -/+	-14.31%	-12.61%	-13.08%	-11.06%	
RATS est.WL_f0	2.909	0.3437	0.966	0.3511	
Bias RATS est.WL_f0	-0.4709	0.3437	-0.4696	0.3511	
RMSE	0.583		0.586		1.006
RATS est.V[est.WL_f0]	0.1012	0.08009	0.0826	0.06100	0.816
n below/above 95% CI	3	377	3	439	
PRB	-13.93%		-13.94%		
PRB -/+	-14.56%	-13.30%	-14.58%	-13.29%	
DENSITY					
True Density	0.0605876	0.015780684	0.0605876	0.015780684	
RATS est.D	0.0555	0.019149	0.0575	0.018970	
Bias RATS est.D	-0.005124	0.011982	-0.003082	0.011618	
RMSE	0.01303		0.01201		0.922
RATS est.V[est.D]	0.0002521	0.00016602	0.0002802	0.00016062	1.112
n below/above 95% CI	2	33	0	18	
PRB	-8.46%		-5.09%		
PRB -/+	-9.682%	-7.231%	-6.276%	-3.899%	
DIST est.D	0.0565	0.020357	0.0587	0.020400	
Bias DIST est.D	-0.004042	0.013539	-0.001880	0.013639	
RMSE	0.01412		0.01376		0.974
DIST est.V[est.D]	0.0002805	0.00019721	0.0003227	0.00019902	1.150
n below/above 95% CI	2	36	1	28	
PRB	-6.67%		-3.10%		
PRB -/+	-8.056%	-5.286%	-4.498%	-1.708%	
RATS est.WL_D	0.05611	0.019197	0.05751	0.018970	
Bias RATS est.WL_D	-0.004477	0.012055	-0.003082	0.011618	
RMSE	0.01285		0.01201		0.935
RATS est.V[est.WL_D]	2.71E-04	1.718E-04	2.80E-04	1.606E-04	1.036
n below/above 95% CI	1	30	0	18	
PRB	-7.39%		-5.09%		
PRB -/+	-8.62%	-6.16%	-6.28%	-3.90%	

B.6 Example CSR Simulations

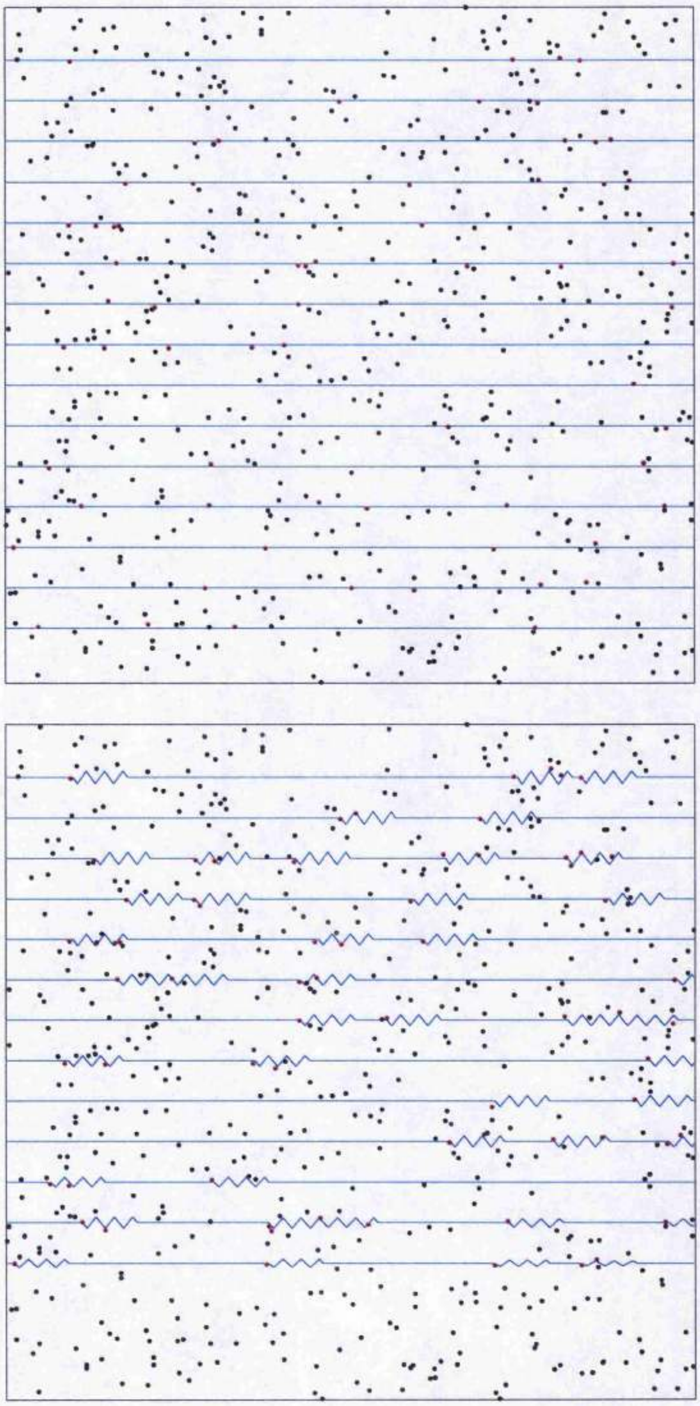


Figure B.1: Example conventional (top) and adaptive (bottom) simulation for the same CSR population. The population size was 600, there were 64 sightings for the conventional survey and 68 for the adaptive survey. Transects are shown as blue lines, the undetected objects as black dots and detected objects red dots.

B.7 Example Clustered Simulations

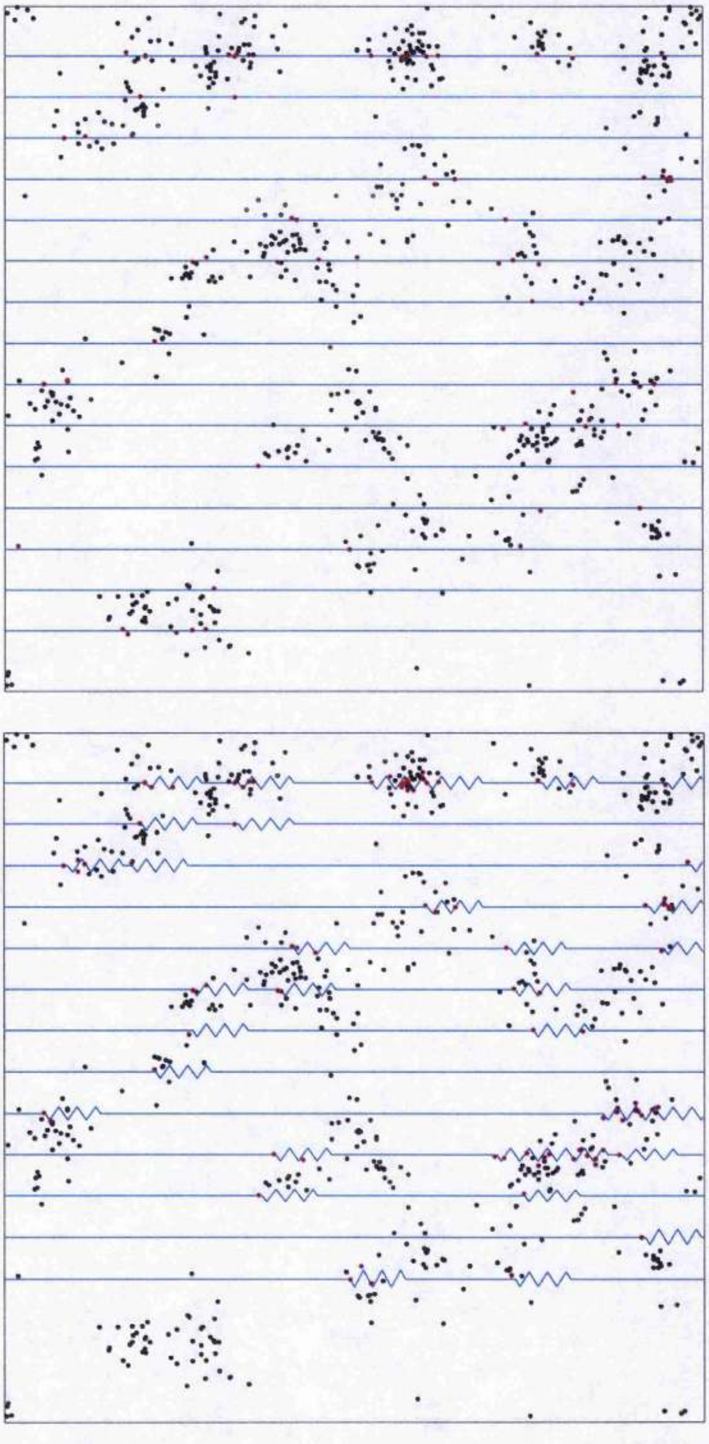
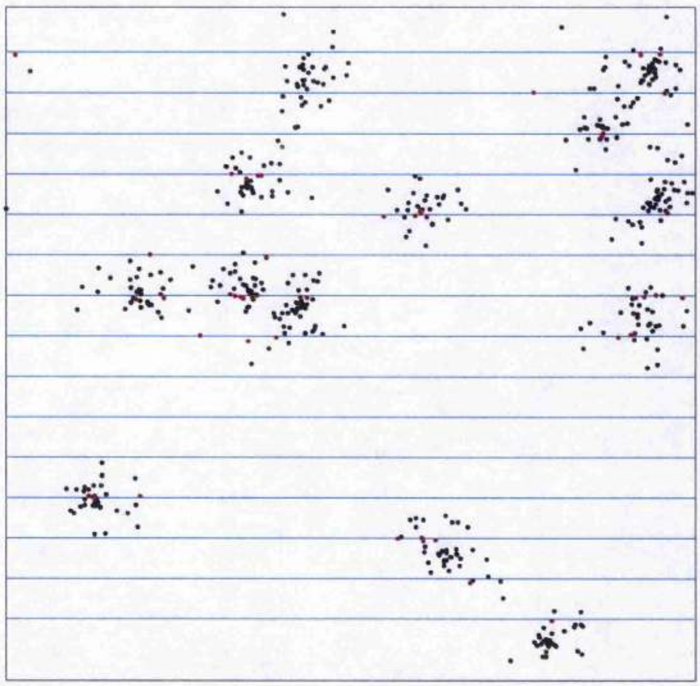


Figure B.2: Example conventional (top) and adaptive (bottom) simulation for the same Clustered population. The population size was 670, there were 70 sightings for the conventional survey and 92 for the adaptive survey. Transects are shown as blue lines, the undetected objects as black dots and detected objects red dots.

B.8 Example Highly Clustered Simulations



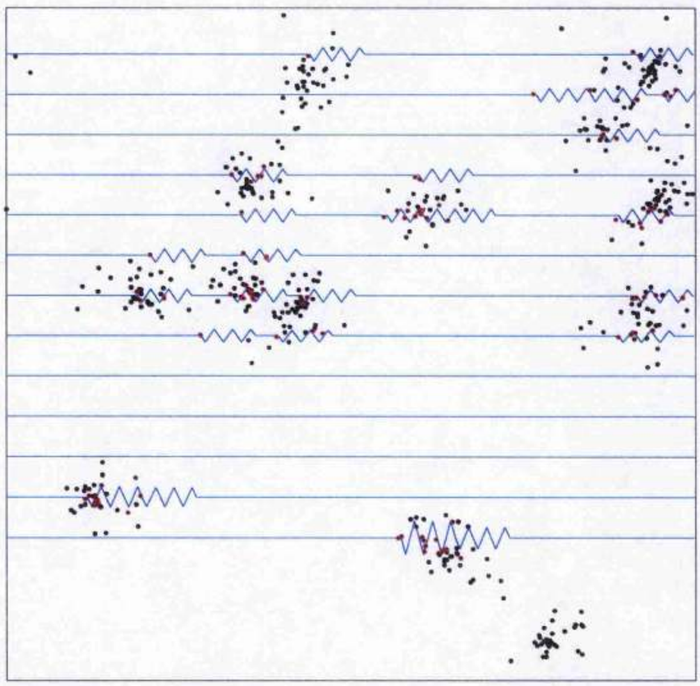


Figure B.3: Example conventional (top) and adaptive simulation (bottom) for the same Highly Clustered population. The population size was 517, there were 52 sightings for the conventional survey and 104 for the adaptive survey. Transects are shown as blue lines, the undetected objects as black dots and detected objects red dots.

Bibliography

Alho, J.M. (1990). Logistic regression in capture-recapture models. *Biometrics* 46, 623-635.

Anderson, D.R., Burnham, K.P., White, G.C. and Otis, D.L. (1983). Density estimation of small-mammal populations using a trapping web and distance sampling methods. *Ecology*, **64**, 674-80.

Borchers, D.L. (1996). *Line transect abundance estimation with uncertain detection on the trackline*. Unpublished PhD dissertation. University of Cape Town. 233pp.

Borchers, D.L., Buckland, S.T., Goedhart, P.W., Clarke, E.D., Hedley, S.L. (1998a). Horvitz-Thompson Estimators for Double-Platform Line Transect Surveys. *Biometrics* **54**, 1221-1237.

Borchers, D.L., Zucchini, W. and Fewster, R.M. (1998b). Mark-Recapture Models for Line Transect Surveys. *Biometrics* **54**, 1207-1220.

Brown, B.M. and Cowling, A. (1998). Clustering and abundance estimation for Neyman-Scott models and line transect surveys. *Biometrika* **85**, 427-438.

Brown, J.A. (1996). The relative efficiency of adaptive cluster sampling for ecological surveys. *Mathematical and Information Sciences Reports, series B*: 96/08, Massey University, New Zealand.

Brown, J.A. (1999). A comparison of two adaptive sampling designs. *Australian and New Zealand Journal of Statistics* **41**, 395-403.

Brown, J.A., and Manly, B. (1998). Restricted adaptive cluster sampling. *Environmental and Ecological Statistics* **5**, 49-63.

Buckland, S.T. (1992). Effects of heterogeneity on estimation of probability of detection on the trackline. *Report of the International Whalling Commission*. **42**, 569-573

Buckland, S.T., Anderson, D.R., Burnham, K.P., and Laake, J.L. (1993). *Distance Sampling: Estimating abundance of biological populations*. London: Chapman & Hall.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. and Thomas, L. (2001). *Introduction to Distance Sampling*. Oxford: Oxford University Press.

Buckland, S.T., Goudie, I.B.J. and Borchers, D.L. (2000). Wildlife Population Assessment: Past Developments and Future Directions. *Biometrics* **56**, 1-12.

Buckland, S.T., Thomas, L., Marques, F.F.C., Strindberg, S., Hedley, S.L., Pollard, J.H., Borchers, D.L. and Burt, M.L. (2002). Distance sampling: recent advances and future directions. In *Quantitative Methods for Current Environmental Issues*. C.W. Anderson, V. Barnett, P. Chatwin and A. El-Shaarawi (eds). London: Springer Verlag.

Burnham, K.P., Anderson, D.R. and Laake, J.L. (1980). Estimation of density from line transect sampling of biological populations. *Wildlife Monographs*, **72**, 1-202.

Christman, M.C. (1997). Efficiency of some sampling designs for spatially clustered populations. *Environmetrics* **8**, 145-166.

Christman, M.C. and Lan, F. (2001). Inverse Adaptive Cluster Sampling. *Biometrics* **57**, 1096-1105.

Christman, M.C. and Pontius, J.S. (2000). Bootstrap confidence intervals for adaptive cluster sampling. *Biometrics* **56**, 503-510.

Cowling, A. (1998). Spatial methods for line transect surveys. *Biometrics* **54**, 828-839.

Cressie, N.A.C. (1993). *Statistics for spatial data revised edition*. USA: Wiley-Interscience.

Diggle, P.J. (1983). *Statistical Analysis of Spatial Point Patterns*. London: Academic Press.

Efron, B. and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap*. London: Chapman and Hall.

Fewster, R.M. and Buckland, S.T. (in preparation). Distance sampling: its relevance to wildlife management. *Journal of the Royal Statistical Society. Series B*.

Fisher, R.A., Thornton, H.G., and Mackenzie, W.A. (1922). The accuracy of the plating method of estimating the density of bacterial populations. *Annals of Applied Biology*, **9**, 325-359.

Garret-Logan, N. and Smith, T. (1997). A hand-held, pen-based computer system for marine mammal sighting surveys. *Marine Mammal Science*. **13(4)**, 694-700.

Hammond, P.S., Benke, H., Berggren, P., Borchers, D.L., Buckland, S.T., Collet, A., Heide-Jorgensen, M.P., Heimlich-Boran, S., Hiby, A.R., Leopold, M.P. and Oien, N. (1995). *Distribution and abundance of the harbour porpoise and other small cetaceans in the North Sea and adjacent waters.* LIFE 92-2/UK/027 Final Report, Commission of the European Union.

Hansen, M.M. and Hurwitz, W.N. (1943). On the theory of sampling from finite populations. *Annals of Mathematical Statistics,* **14,** 333-362.

Hedley, S.L. (2000). *Modelling heterogeneity in cetacean surveys*. Unpublished PhD dissertation. University of St. Andrews. 132pp.

Hedley, S.L., Buckland, S.T. and Borchers, D.L. (1997). Spatial modelling from line transect data. *Journal of Cetacean Research and Management*. **1**, 255-264.

Hiby, A.R. (1982). Using average number of whales in view to estimate population density. *Report of the International Whaling Commission,* **32**, 562-5.

Hiby, A.R. (1985). An approach to estimating population densities of great whales from sighting surveys. *IMA Journal of Mathematics Applied in Medicine and Biology*, **2**, 201-20.

Hiby, A.R. and Hammond, P.S. (1989). Survey techniques for estimating current abundance and monitoring trends in abundance of cetaceans. In *The Comprehensive Assessment of Whale Stocks: the early years*, Donovan, G.P. (ed.), pp. 47-80. Cambridge: International Whaling Commission.

Horvitz, D.G. and Thompson, D.G. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* **47**, 663-685.

Huggins, R.M. (1989). On the statistical analysis of capture experiments. *Biometrika*. **76**, 133-40.

Laake, J.L., Buckland, S.T., Anderson, D.R. and Burnham, K.P. (1994). *DISTANCE User's Guide*. Fort Collins, Colorado: Colorado Cooperative Fish and Wildlife Research Unit, Colorado State University.

Laake, J.L., Buckland, S.T., Anderson, D.R. and Burnham, K.P. (1996). *DISTANCE 2.2*. Colorado Cooperative Fish and Wildlife Research Unit, Colorado State University, Fort Collins, Colorado. Available: http://www.ruwpa.st-and.ac.uk/distance/

Levene, H. (1960). In *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, I. Olkin et al. eds., Stanford University Press, pp. 278-292.

Lindberg, L and Rexstad, E. (2002). Capture-recapture sampling designs. In *Encyclopedia of Environmetrics*, El-Shaarawi, A.H. and Piegorsch, W.W. (eds.), Volume 1, 251-262. Chichester: John Wiley and Sons, Ltd.

Manly, B.F.J. (1997). *RT, A Program for Randomisation Testing, Version 2.0*. Oregon: Centre for Applications of Statistics and Mathematics, University of Oregon.

Manly, B.F.J. (1997). *Randomisation, Bootstrap and Monte Carlo Methods in Biology*. 2nd Edition. London: Chapman and Hall.

Marques, F.F.C. (2001). Estimating wildlife distribution and abundance from line transect surveys conducted from platforms of opportunity. Unpublished PhD dissertation. University of St. Andrews. 152pp.

Marques, F.C., Buckland, S.T., Goffin, D., Dixon, C.E., Borchers, D.L., Mayle, B.A. and Peace, A.J. (2001). Estimating deer abundance from line transect surveys of dung: sika deer in southern Scotland. *Journal of Applied Ecology*, **38**, 349-363.

Palka, D.L. (1993). *Estimating density of animals when assumptions of line-transect surveys are violated*. Unpublished PhD dissertation. University of California, San Diego. 169pp.

Palka, D.L. (Palka, D. (1995). Abundance estimate of Gulf of Maine harbor porpoise. *Report of the International Whaling Commission* Special issue **16**, 27-50.

Palka, D. (1996). Update on abundance of Gulf of Maine/Bay of Fundy harbor porpoises. *Northeast Fisheries Science Center Ref. Doc. 96-04*. Woods Hole, Maine: NOAA/NMFS/NEFSC.

Palka, D.L. and Pollard, J.H. (1999). Adaptive line transect survey for harbor porpoises. In *Marine Mammal Survey and Assesssment Methods*, G.W. Garner, S.C. Amstrup, J.L. Laake, B.F.J. Manly, L.L. McDonald and D.G. Robertson (eds.), pp. 3-11. Rotterdam: A.A. Balkema.

Parmenter, R.P., Yates, T.L., Anderson, D.R., Burnham, K.P., Dunnum, J.L., Franklin, A.B., Friggens, M.T., Lubow, B.C., Miller, M, Olson, G.S., Parmenter, C.A., Pollard, J.H., Rexstad, E., Shenk, T.M., Stanley, T.R. and White, G.C. (2002). Small mammal density estimation: A field comparison of grid-based versus web-based density estimators. *Ecological Monographs* (in press).

Pollard, J.H. and Buckland, S.T. (1997). A strategy for adaptive sampling in shipboard line transect surveys. *Report of the International Whaling Commission.* **47**, 921-931.

Pollard, J.H., Palka, D.L. and Buckland, S.T. (2002). Adaptive distance sampling. *Biometrics* (in press).

Ramsey, F.L., Wildman, V. and Engbring, J. (1987). Covariate adjustments to effective area in variable-area wildlife surveys. *Biometrics* **43**, 1-11.

Salehi, M.M. and Seber, G.A.F. (1997). Two-stage adaptive cluster sampling. *Biometrics* **53**, 959-970.

Särndal, C.E, Swensson, B. and Wretman, J. (1992). *Model Assisted Survey Sampling*. New York: Springer-Verlag.

Seber, G.A.F. (1982). *The Estimation of Animal Abundance and related parameters*. 2nd Edition. London: Edward Arnold.

Smith, T.D., Palka, D. and Bisack, K. (1993). Biological significance of bycatch of harbor porpoise in the Gulf of Maine demersal gillnet fishery. *Northeast Fisheries Science Center Ref. Doc. 93-23*. Woods Hole, Maine: NOAA/NMFS/NEFSC.

Steinhorst, R.K. and Samuel, M.D. (1989). Sightability adjustment methods for aerial surveys of wildlife populations. *Biometrics* **45**, 415-425.

Strindberg, S. (2001). *Optimised Automated Survey Design in Wildlife Population Assessment*. Unpublished PhD dissertation. University of St. Andrews. 356pp.

Statistical Sciences Inc. (1997). *S-Plus user's manual, version 4.5 for Windows.* Seattle: Statistical Sciences Inc.

Thomas, L., Laake, J.L., Strindberg, S., Marques, F.F.C., Borchers, D.L., Buckland, S.T., Anderson, D.R., Burnham, K.P., Hedley, S.L. and Pollard, J.H. (2002). *Distance 4.0. Beta 6.* Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. Available: http://www.ruwpa.st-and.ac.uk/distance/

Thomas, L., Buckland, S.T., Burnham, K.P., Anderson, D.R., Laake, J.L., Borchers, D.L. and Strindberg, S. (2002). Distance sampling. In *Encyclopedia of Environmetrics*, El-Shaarawi, A.H. and Piegorsch, W.W. (eds), Volume 1, 544-552. Chichester: John Wiley and Sons, Ltd.

Thompson, S.K. (1990). Adaptive cluster sampling. *Journal of the American Statistical Association* **85**, 1050-1059.

Thompson, S.K. (1991). Stratified adaptive cluster sampling. *Biometrika* **78**, 389-397.

Thompson, S.K. (1992). *Sampling*. New York: Wiley Interscience.

Thompson, S.K. (1994). Factors influencing the efficiency of adaptive cluster sampling. *Technical Report* 94-0301, Center for Statistical Ecology and Environmental Statistics, Department of Statistics, Pennsylvania State University, USA.

Thompson, S.K. and Ramsey F.L. (1987). Detectability Functions in Observing Spatial Point Processes. *Biometrics* **43**, 355-362.

Thompson, S.K., Ramsey F.L. and Seber, G.A.F. (1992). An Adaptive Procedure for Sampling Animal Populations. *Biometrics* **48**, 1195-1199.

Thompson, S.K. and Seber, A.F. (1994). Detectability in Conventional and Adaptive Sampling. *Biometrics* **50**, 712-724.

Thompson, S.K. and Seber, A.F. (1996). *Adaptive Sampling*. New York: Wiley Interscience.

Yates, F. and Grundy, P.M. (1953). Selection Without Replacement from Within Strata with Probability Proportional to Size. *Journal of the Royal Statistical Society. Series B* **15**, 253-261.